# Contents

# Lectures on General Relativity

Valeria Ferrari

# 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 foundamental 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, that we shall investigate in this course disclose interesting and fascinating new phenomena. The gravitational collapse and the formation of black holes, the existence of gravitational waves, the Big Bang of cosmological theories are, at least conceptually, in the common background of modern physicists.

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 development of the mathematical tools and techniques that are essential to describe the theory and the physical consequences. The first lecture 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 need to introduce the concept of manifold, metric, affine connections and other geometrical objects.

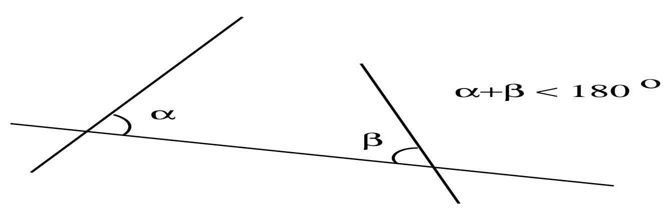
3) What is the role played by the equivalence principle in all that.

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

But first of all, since as we have already anticipated 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 Euclide's postulates. Among them the fifth has been the object of a millennary dispute: for over 2000 years geometers tried to show, without succeeeding, that the fifth postulate is a consequence of the other four. The postulate states
*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 $180^0$ , the two lines will meet at some point on the side of the internal angles.*




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 Euclide'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

$$x_1^2 + x_2^2 < 1, \tag{1.1}$$

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

$$d(x, X) = a \cosh^{-1} \left[ \frac{1 - x_1 X_1 - x_2 X_2}{\sqrt{1 - x_1^2 - x_2^2}\sqrt{1 - X_1^2 - X_2^2}} \right], \tag{1.2}$$

where $a$ is a lenghtscale. This space is infinite, because $d(x, X) \to \infty$ when $X_1^2 + X_2^2 \to 1$. The logical independence of Euclide'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 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 or 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 realtionships between points on the surface.

However, since a cilinder 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 that stay on them 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 $(x_1, x_2)$ that *does* cover the space. How do we express the *distance* between two points? If we explicitely evaluate $d\xi_1$ and $d\xi_2$ in terms of the new coordinates we find

$$\begin{aligned} \xi_1 &= \xi_1(x_1, x_2) \qquad \to d\xi_1 = \frac{\partial \xi_1}{\partial x_1} dx_1 + \frac{\partial \xi_1}{\partial x_2} dx_2 \\ \xi_2 &= \xi_2(x_1, x_2) \qquad \to d\xi_2 = \frac{\partial \xi_2}{\partial x_1} dx_1 + \frac{\partial \xi_2}{\partial x_2} dx_2 \end{aligned} \tag{1.4}$$

$$\begin{aligned} ds^2 &= \left[ \left( \frac{\partial \xi_1}{\partial x_1} \right)^2 + \left( \frac{\partial \xi_2}{\partial x_1} \right)^2 \right] dx_1^2 + \left[ \left( \frac{\partial \xi_1}{\partial x_2} \right)^2 + \left( \frac{\partial \xi_2}{\partial x_2} \right)^2 \right] dx_2^2 \\ &+ 2 \left[ \left( \frac{\partial \xi_1}{\partial x_1} \right) \left( \frac{\partial \xi_1}{\partial x_2} \right) + \left( \frac{\partial \xi_2}{\partial x_1} \right) \left( \frac{\partial \xi_2}{\partial x_2} \right) \right] dx_1 dx_2 \\ &= g_{11} dx_1^2 + g_{22} dx_2^2 + 2 g_{12} dx_1 dx_2 = g_{\alpha\beta} dx^\alpha dx^\beta. \end{aligned} \tag{1.5}$$

In the last line of eq. (1.6) we have defined the following quantities:

$$\begin{aligned} g_{11} &= \left[ \left( \frac{\partial \xi_1}{\partial x_1} \right)^2 + \left( \frac{\partial \xi_2}{\partial x_1} \right)^2 \right] \\ g_{22} &= \left[ \left( \frac{\partial \xi_1}{\partial x_2} \right)^2 + \left( \frac{\partial \xi_2}{\partial x_2} \right)^2 \right] \\ g_{12} &= \left[ \left( \frac{\partial \xi_1}{\partial x_1} \right) \left( \frac{\partial \xi_1}{\partial x_2} \right) + \left( \frac{\partial \xi_2}{\partial x_1} \right) \left( \frac{\partial \xi_2}{\partial x_2} \right) \right], \end{aligned} \tag{1.6}$$

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 preceeding 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 CONVENCTION

In writing the last line of eq. (1.6) we have adopted the convenction 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 a summation is implied. For example, if the index $\alpha$ takes the values 1 and 2

$$x_\alpha X^\alpha = \sum_{i=1}^{2} x_i X^i = x_1 X^1 + x_2 X^2 \tag{1.7}$$

We shall adopt this convenction 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

$$\xi_1 = r\cos\theta \qquad \rightarrow d\xi_1 = \cos\theta dr - r\sin\theta d\theta \tag{1.8}$$

$$\xi_2 = r\sin\theta \qquad \rightarrow d\xi_2 = \sin\theta dr + r\cos\theta d\theta$$

$$\tag{1.9}$$

$$ds^2 = (d\xi^1)^2 + (d\xi^2)^2 = dr^2 + r^2 d\theta^2, \tag{1.10}$$

and therefore

$$g_{11} = 1, \qquad g_{22} = r^2, \qquad g_{12} = 0. \tag{1.11}$$

***********************************************************

## 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 co-ordinate transformation. Let us suppose 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 $(x_1', x_2')$ . We have seen in section 1 that, for example, the component $g_{11}$ is defined as

$$g_{11} = [(\frac{\partial \xi_1}{\partial x_1})^2 + (\frac{\partial \xi_2}{\partial x_1})^2], \tag{1.12}$$

where $(\xi_1, \xi_2)$ are the coordinates of the locally euclidean reference frame, and $(x_1, x_2)$ two arbitrary new coordinates. If we now put $x_1 = x_1(x_1', x_2')$ , and $x_2 = x_2(x_1', x_2')$ , the metric tensor in the new coordinate frame $(x_1', x_2')$ will be

$$\begin{aligned} g_{11}' &= [(\frac{\partial \xi_1}{\partial x_1'})^2 + (\frac{\partial \xi_2}{\partial x_1'})^2] \qquad (1.13) \\ &= [(\frac{\partial \xi_1}{\partial x_1}\frac{\partial x_1}{\partial x_1'} + \frac{\partial \xi_1}{\partial x_2}\frac{\partial x_2}{\partial x_1'})^2 + [(\frac{\partial \xi_2}{\partial x_1}\frac{\partial x_1}{\partial x_1'} + \frac{\partial \xi_2}{\partial x_2}\frac{\partial x_2}{\partial x_1'})^2 \\ &= [(\frac{\partial \xi_1}{\partial x_1})^2 + (\frac{\partial \xi_2}{\partial x_1})^2](\frac{\partial x_1}{\partial x_1'})^2 + [(\frac{\partial \xi_1}{\partial x_2})^2 + (\frac{\partial \xi_2}{\partial x_2})^2](\frac{\partial x_2}{\partial x_2'})^2 \end{aligned}$$

$$
\begin{aligned}
&+ \quad 2(\frac{\partial \xi_1}{\partial x_1}\frac{\partial \xi_1}{\partial x_2} + \frac{\partial \xi_2}{\partial x_1}\frac{\partial \xi_2}{\partial x_2})(\frac{\partial x_1}{\partial x'_1}\frac{\partial x_2}{\partial x'_1}) \\
&= \quad g_{11}(\frac{\partial x_1}{\partial x'_1})^2 + g_{22}(\frac{\partial x_2}{\partial x'_2})^2 + 2g_{12}(\frac{\partial x_1}{\partial x'_1}\frac{\partial x_2}{\partial x'_1}).
\end{aligned}
$$

In general we can write

$$g'_{\alpha\beta} = g_{\mu\nu}\frac{\partial x_\mu}{\partial x'^\alpha}\frac{\partial x_\nu}{\partial x'^\beta} \tag{1.14}$$

This is the manner in which tensors transfor under an arbitrary coordinate transformation (This point will be illustrated in more detail in following lectures).

************************************************************

Thus, given a space in which the *distance* can be expressed in terms of Pythagoras' law, if we make an arbitrary coordinate transformation the knowledge of $g_{\mu\nu}$ allows us to express the *distance* in the new reference system. The converse is also true: given a space in which

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta, \tag{1.15}$$

if this space belongs to the class defined by Gauss, at any given point it is always possible to choose a locally euclidean coordinate system $(\xi_\alpha)$ such that

$$ds^2 = d\xi_1^2 + d\xi_2^2. \tag{1.16}$$

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

**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 that depends on the inner properties of the surface, but does not depend on the particular coordinate system we may 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

$$\begin{aligned}
k(x_1,x_2) = \frac{1}{2g}\left[2\frac{\partial^2 g_{12}}{\partial x_1 \partial x_2} - \frac{\partial^2 g_{11}}{\partial x_2^2} - \frac{\partial^2 g_{22}}{\partial x_1^2}\right] \\
-\frac{g_{22}}{4g^2}\left[\left(\frac{\partial g_{11}}{\partial x_1}\right)\left(2\frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1}\right) - \left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right] \\
+\frac{g_{12}}{4g^2}\left[\left(\frac{\partial g_{11}}{\partial x_1}\right)\left(\frac{\partial g_{22}}{\partial x_2}\right) - 2\left(\frac{\partial g_{11}}{\partial x_2}\right)\left(\frac{\partial g_{22}}{\partial x_1}\right)\right. \\
\left. + \left(2\frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2}\right)\left(2\frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1}\right)\right] \\
-\frac{g_{11}}{4g^2}\left[\left(\frac{\partial g_{22}}{\partial x_2}\right)\left(2\frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2}\right) - \left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]
\end{aligned} \tag{1.17}$$

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

$$g = g_{11}g_{22} - g_{12}^2. \tag{1.18}$$

For example, given a spherical surface, no matter how do we choose the coordinates, we shall always find that

$$k = \frac{1}{a^2}, \tag{1.19}$$

(try for example a sphere in polar coordinates $ds^2 = a^2 d\theta^2 + a^2 \sin^2\theta d\varphi^2$, where $a$ is the radius), or, for the Gauss-Bolyai-Lobachewski geometry where

$$g_{11} = \frac{a^2(1-x_2^2)}{(1-x_1^2-x_2^2)^2}, \qquad g_{22} = \frac{a^2(1-x_1^2)}{(1-x_1^2-x_2^2)^2}, \qquad g_{12} = \frac{a^2 x_1 x_2}{(1-x_1^2-x_2^2)^2}, \tag{1.20}$$

we shall always find

$$k = -\frac{1}{a^2}, \tag{1.21}$$

or, if the space is flat, $k = 0$ . 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 shown that there exists a scalar quantity expressing the inner properties of a surface that is a function of $g_{\alpha\beta}$ and its first and second derivatives, i.e. the gaussian curvature, that it is invariant under a coordinate transformation.

These results can be extended to an arbitrary D-dimensional space. In particular, as we shall understand better 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.

$$ds^2 = -(d\xi^0)^2 + (d\xi^1)^2 + (d\xi^2)^2 + (d\xi^3)^2. \tag{1.22}$$

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 symmetic, 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

$$C = \frac{D(D+1)}{2} - D = \frac{D(D-1)}{2}. \tag{1.23}$$

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

## 1.4 The newtonian theory

In this section we shall understand why the newtonian theory of gravity became unappropriate to correctly deescribe the gravitational field. The newtonian theory of gravity was published in 1685 in the "Philosophiae Naturalis Principia Mathematica". They contain an incredible variety of fundamental results, and among them, the corner stones of classical physics:

1) *Newton's law*

$$\vec{F} = m_I \vec{a}, \tag{1.24}$$

2) *Newton's law of gravitation*

$$\vec{F}_G = m_G \vec{g}, \tag{1.25}$$

where

$$\vec{g} = \frac{G \sum_i M_{Gi}(\vec{r} - \vec{r}_i)}{|\vec{r} - \vec{r}_i|^3} \tag{1.26}$$

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

$$\vec{a} = \left(\frac{m_G}{m_I}\right) \vec{g}. \tag{1.27}$$

If $\frac{m_G}{m_I}$ is a constant independent on the body, the acceleration of falling bodies is the same for every body and independent on their mass. Galileo had already experimentally discovered that this is true indeed. Newton itself tested the equivalence principle studying the motion of pendulum of different composition and equal lenght, and he found 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. 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.*

I do not know if you have ever appreciated the big conceptual step that is implied in this sentence.

Since then a variety of experiments confirmed this crucial result. Among them Eotvos experiment in 1889 (accuracy of 1 part in $10^9$), and Dicke experiment in 1964 (1 part in $10^{11}$), and Braginsky in 1972 (1 part in $10^{12}$). All the 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 describe briefly the reasons of his success, that remained untouched for more than 200 years.

The monumental construction of the Principia is based on the newtonian law of gravitation. The theory of lunar motion and tides, the description of the planetary motion around the sun are the most elegant and accomplished description of these phenomena. It is interesting to note that in the Principia, Newton does not write equations: he simply describes in words what we are now able to compute using his equations and the mathematical instrument of infinitesimal calculus he invented and developed.

After Newton, the law of gravitation was used to investigate in more detail the solar system, and its application in the study of the perturbations of the orbit of Uranus led, in 1846, Adams (England) and Le Verrier (France) to the prediction of the existence of Neptune. A few years later the discovery of Neptun was a triumph of Newton's theory.

However, already in 1845 Le Verrier had observed anomalies in the motion of Mercury. He found that the precession in the perihelium of $35''/100 years$ exceeded the value due to the perturbation introduced by the other planets according to 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 of the exceedingly high precession of the perihelium of Mercury has a relativistic origin.

In any event, we can say that the newtonian theory worked remarkably fine 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 phylosophical aspect of the theory. The equations of newtonian mechanics are invariant under Galileo's transformation

$$
\begin{aligned}
\vec{x}' &= R\vec{x} + \vec{v}t + \vec{d} \\
t' &= t + \tau
\end{aligned}
\tag{1.28}
$$

where $R$ is the orthogonal matrix expressing three-dimensional rotations (its element depend on the three Euler angles), $\vec{v}$ is the relative velocity of the two references, and $\vec{d}$ the 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** in which the laws of mechanics hold. What then determines which reference 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. Thus inertial frames are those in uniform relative motion with respect to the absolute space. 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 the motion of the water in the rotating vessel is due to the particular distribution of masses in the universe rather than to the relative motion of the vessel with respect to absolute space. An interesting discussion on the subject can be found in Weinberg pg. 15-16-17, and we shall not discuss it further here. For our present purposes it is enough to realise that another element of 'dissatisfaction' of the newtonian theory is due to the need of introducing an absolute space.

The problems that 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 ninentheenth century. The big storm approaches with the formulation of the theory of electrodynamics presented by Maxwell in 1864. Maxwell's equations state that the velocity of light must be a universal constant. It was soon understood that these equations are not invariant under Galileo's transformations, because according to them if the velocity of light is $c$ in a given coordinate system, it cannot be $c$ in a second reference moving with respect to the first according to eqs. (1.28). To justify this discrepancy, Maxwell formulated the hypothesis that light does not really propagate in vacuum: electromagnetic waves are carried by a medium, the *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 direction of the earth's orbital motion, and transverse to it. What could it be the explanation of this fact? Either the earth could be in quiet with respect to the ether, but this hypothesis was totally unsatisfactory: it would have been a coming back to an antropocentric picture of the world. Or simply the ether did not exist and one has to accept the fact that the speed of light is the same in any direction. 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, who showed that Galileo's transformations had 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

$$L^0_0 = \gamma, \quad L^0{}_j = L^j_0 = \frac{\gamma}{c} v_j, \qquad L^i{}_j = \delta^i{}_j + \frac{\gamma - 1}{v^2} v_i v_j. \quad i, j = 1, 3 \tag{1.30}$$

and $v^i$ are the components of the velocity of the boost.

As it was immediately realised, however, while Maxwell's equations were invariant with respect to Lorentz transformations, Newton's equations were not, and it started to become clear that 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 set of non relativistic particles subjected to arbitrary forces $\vec{F}$ and moving in a gravitational field that is **constant** . According to newtonian mechanics, the equation of motion will be

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

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

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

In this new reference eq. (1.31) becomes

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

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

$$(m_I)_n \frac{d^2 \vec{x}'_n}{dt^2} = \sum_k \vec{F}(\vec{x}'_n - \vec{x}'_k), \tag{1.34}$$

Compare eq. (1.31) and eq. (1.34). We understand 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 see the gravitational field. **This result follows from the equivalence, experimentally tested, of the inertial and the 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 we 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 not subjected to any other force. 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 unifom everywhere the coordinate transformation (1.32) defines a locally inertial frame that will cover the whole spacetime. If this is not the case we can set up a locally inertial frame only in the neighbourhood 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, as we shall see soon, to describe the effects of gravity in terms of curved geometry. Equation (1.34) has been derived in the framework of newtonian theory, but we know how to generalize it in Special Relativity: time will be replaced by proper time, 3-vectors by fourvectors, 3-forces by four-forces, but what we said about the possibility of eliminating the gravitational field by a suitable choice of the reference still holds. So 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.*

Now it is clear that the preceeding formulation of the equivalence principle resembles very much to the axiom that Gauss chose as a basis for non-euclidean geometries: 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 the Minkowsky expression

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = -(d\xi^0)^2 + (d\xi^1)^2 + (d\xi^2)^2 + (d\xi^3)^2. \quad (1.35)$$

We therefore expect that the equations of gravity will look very similar to those of Riemaniann 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 eq. (1.6) 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, by virtue 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 equation of motion for 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, for example the one in free fall with the particle. For the principle of Equivalence, in this frame the *distance* will be

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \eta_{\mu\nu} d\xi^\mu d\xi^\nu, \tag{1.36}$$

where $\eta_{\mu\nu} = diag(-1, 1, 1, 1)$ is the metric tensor of the flat Minkowskian spacetime. If $\tau$ is the time coordinate in this frame, i.e. $\tau$ is the proper time of the particle, the equation of motion, for what we said before, are

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

We now change to a frame where the coordinates are labelled $x^\alpha = x^\alpha(\xi^\alpha)$ . Thus we assign a law of transformation that allows us 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 fram the *distance* is

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

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. (1.7)). In the new frame the equation of motion of the particle (1.37) becomes:

$$\frac{d^2x^\alpha}{d\tau^2} + \left[\frac{\partial x^\alpha}{\partial\xi^\lambda}\frac{\partial^2\xi^\lambda}{\partial x^\mu\partial x^\nu}\right]\left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\right] = 0. \tag{1.40}$$

(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

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

The quantities (1.41) are called the **affine connections**, or the **Christoffel symbols**, whose properties we shall investigate in a following lecture. Equation (1.42) is the **geodesic equation**, 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 virtue of 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 shall see 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 reference where they vanish), i.e. the derivative 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 explicitely, and we shall show that this is indeed the case in a next lecture . The result will be

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

**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 $\frac{\partial \xi^{\beta}}{\partial x^{\lambda}}$

$$\frac{\partial \xi^{\beta}}{\partial x^{\lambda}}\Gamma^{\lambda}_{\mu\nu} = \frac{\partial \xi^{\beta}}{\partial x^{\lambda}}\frac{\partial x^{\lambda}}{\partial \xi^{\alpha}}\frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu}\partial x^{\nu}} = \delta^{\beta}_{\alpha}\frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu}\partial x^{\nu}} = \frac{\partial^2 \xi^{\beta}}{\partial x^{\mu}\partial x^{\nu}}, \tag{1.45}$$

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 $X$

$$\xi^{\beta}(x) = \xi^{\beta}(X) + [\frac{\partial \xi^{\beta}(x)}{\partial x^{\lambda}}\Gamma^{\lambda}_{\mu\nu}]_{x=X}(x^{\mu} - X^{\mu})(x^{\nu} - X^{\nu}) + ... = a^{\beta} + b^{\beta}_{\lambda}\Gamma^{\lambda}_{\mu\nu}(x^{\mu} - X^{\mu})(x^{\nu} - X^{\nu}) + ... \tag{1.47}$$

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

$$g_{\mu\nu}(X) = \eta_{\alpha\beta}\frac{\partial \xi^{\alpha}(x)}{\partial x^{\mu}}|_{x=X}\frac{\partial \xi^{\beta}(x)}{\partial x^{\nu}}|_{x=X} = \eta_{\alpha\beta}b^{\beta}_{\mu}b^{\beta}_{\nu}, \tag{1.48}$$

and from this equation we compute $b^{\beta}_{\mu}$. Thus, given $g_{\mu\nu}$ and $\Gamma^{\alpha}_{\mu\nu}$ 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

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0, \qquad (A1)$$

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

$$\xi^\alpha = \xi^\alpha(x^\gamma), \qquad \rightarrow \qquad \frac{d\xi^\alpha}{d\tau} = \frac{\partial\xi^\alpha}{\partial x^\gamma}\frac{dx^\gamma}{d\tau}, \qquad (A2)$$

eq. (A1) becomes

$$\frac{d}{d\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\gamma}\frac{dx^\gamma}{d\tau}\right) = \frac{d^2x^\gamma}{d\tau^2}\frac{\partial\xi^\alpha}{\partial x^\gamma} + \frac{\partial^2\xi^\alpha}{\partial x^\beta\partial x^\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0. \quad (A3)$$

Multiply eq. (A3) by $\frac{\partial x^\sigma}{\partial\xi^\alpha}$ 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 $\sigma = \gamma$ 0 otherwise), we find

$$\frac{d^2x^\gamma}{d\tau^2}\delta^\sigma_\gamma + \frac{\partial x^\sigma}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\beta\partial x^\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0, \quad (A4)$$

which finally becomes

$$\frac{d^2x^\sigma}{d\tau^2} + [\frac{\partial x^\sigma}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\beta\partial x^\gamma}]\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0, \quad (A5)$$

which is eq. (1.40).

## 1.10 Appendix 1B

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

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

let us consider the Lorentz transformation

$$\xi^i = L^i{}_j \xi^j, \qquad (B2)$$

where

$$L^i_j = \delta^i_j + v^i v_j \frac{\gamma - 1}{v^2}, \qquad L^0_j = \frac{\gamma v_j}{c}, \qquad L^0_0 = \gamma, \qquad \gamma = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}. \qquad (B3)$$

The *distance* will now be

$$ds^2 = \eta_{\mu\nu} d\xi^\mu d\xi^\nu = \eta_{\mu\nu} \frac{\partial \xi^\mu}{\partial \xi^{i\prime}} \frac{\partial \xi^\nu}{\partial \xi^{j\prime}} d\xi^{i\prime} d\xi^{j\prime}.; \qquad (B4)$$

Since

$$\frac{\partial \xi^\mu}{\partial \xi^{i\prime}} = L^\mu{}_\beta \delta^\beta{}_{i\prime} = L^\mu{}_{i\prime}, \qquad (B5)$$

it follows that

$$ds^2 = \eta_{\mu\nu} L^\mu{}_{i\prime} L^\nu{}_{j\prime} d\xi^{i\prime} d\xi^{j\prime}. \qquad (B6)$$

$L^i_j$ are constant, consequently we can always rescale the new coordinates in such a way that

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

and the new reference fram is again locally inertial.

# Chapter 2

In chapter I 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 tensor, 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

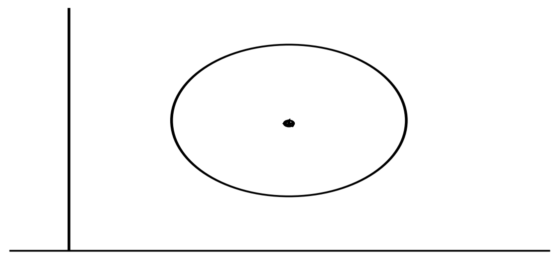
In general relativity we shall deal with *topological spaces.* The word topology has two distinct meanings: local topology (to which we are mainly interested), and global topology, which involves the study of the large scale features of a space, such those that distinguish a sphere from a cone. In order to define a topological space we need some preliminary definitions.

First of all, be $\mathbf{R}^n$ the n-dimensional space of vector algebra: a point in $\mathbf{R}^n$ is a sequence of real numbers $(x^1, x^2, ...x^n)$, also called an n-tuple of real numbers. Intuitively we have an idea that this is a **continuum space**, namely that there are points of $\mathbf{R}^n$ arbitrarily close to any given point, that the line joining two points can be subdivided into arbitrarily many pieces which also join points of $\mathbf{R}^n$. In other words, there are no holes in our space (a non continuous space is, for example, a lattice).

*Open Sets.* Given a point $y = (y^1, y^2, ...y^n)$, an open set is the collection of points $x$ such that

$$|x - y| \equiv \sqrt{\sum_{i=1}^{n}(x^i - y^i)^2} < r, \tag{2.1}$$

where $r$ ís a real number. (This is sometimes called an 'open ball').

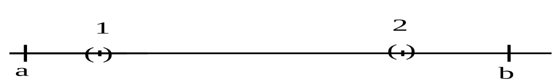


An open set includes all points in the interior of the "ball", *but not on the boundary.*

A set of points **S** is open if every point $x \in \mathbf{S}$ has a neighborhood entirely contained in **S**. Thus if in the definition of open sets we had included the boundary

$$|x - y| \leq r, \tag{2.2}$$

there would exist points in **S** (those on the boundary) that do not satisfy the previous property. The idea that a line connecting two points can be indefinitely subdivided, is related to the so-called *Hausdorff property* of $\mathbf{R}^n$, i.e.: any two points of $\mathbf{R}^n$ have neighborhoods which do not intersect



We can now define open sets of $\mathbf{R}^n$ which satisfy the following properties

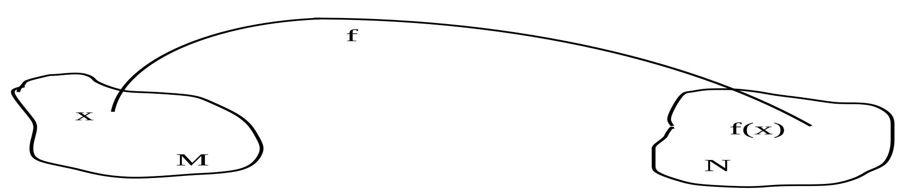
1) if $\mathbf{O}_1$ and $\mathbf{O}_2$ are open sets, so it is their intersection.

2) the union of any collection (possibly infinite in number) of open sets is open.

A topological space is a collection of points that satisfies (1) and (2).

## 2.2 Mapping

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



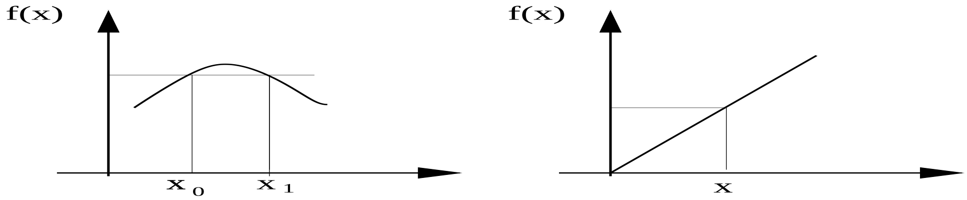

$\mathbf{M}$ and $\mathbf{N}$ need not to be different. For example, the simplest maps are ordinary real-valued functions on $\mathbf{R}$

$$\text{EXAMPLE} \qquad y = x^3, \qquad x \in \mathbf{R}, \qquad \text{and} \qquad y \in \mathbf{R}. \tag{2.3}$$

In this case $\mathbf{M}$ and $\mathbf{N}$ coincide.

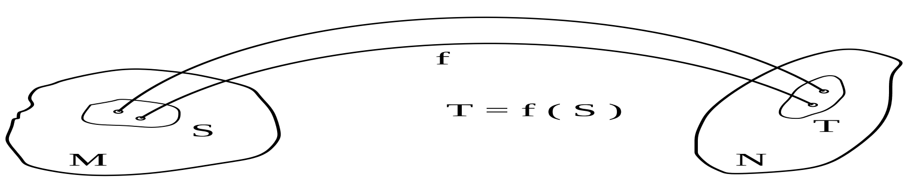
A map gives a unique $f(x)$ for every x, but not necessarily a unique x for every $f(x)$.

EXAMPLE




map **many to one** map **one to one**

If $f$ maps **M** to **N** then for any set **S** in **M** we have an *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**

$$\mathbf{S} = f^{-1}(\mathbf{T}). \tag{2.4}$$

Inverse mapping is possible only in the case of one-to-one mapping. The statement "$f$ maps $\mathbf{M}$ to $\mathbf{N}$" is indicated as

$$f : \mathbf{M} \to \mathbf{N}. \tag{2.5}$$

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

$$f : x \quad | \to y \tag{2.6}$$

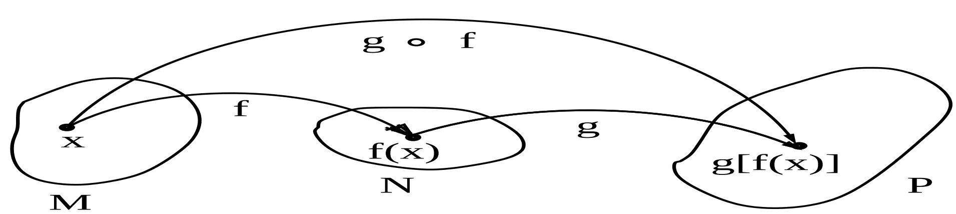
the image of a point x is $f(x)$.

## 2.3 Composition of maps

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

$$g \circ f : \mathbf{M} \to \mathbf{P}. \tag{2.7}$$

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

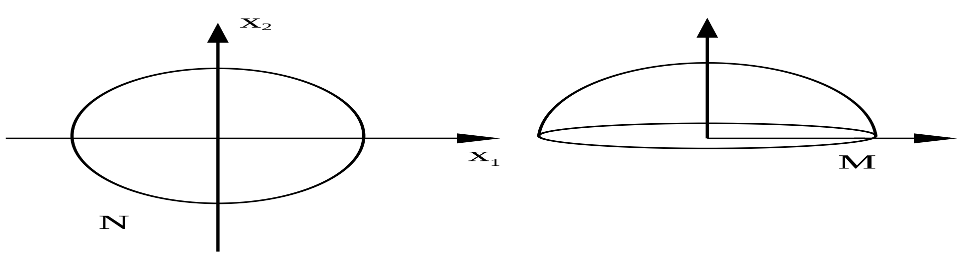


$$
\begin{aligned}
\text{EXAMPLE} \qquad f: \quad x \quad &\mid\rightarrow y \qquad\qquad y = x^3 \\
g: \quad y \quad &\mid\rightarrow z \qquad\qquad z = y^2 \\
g \circ f: \quad x \quad &\mid\rightarrow z \qquad\qquad z = x^6
\end{aligned}
\tag{2.8}
$$

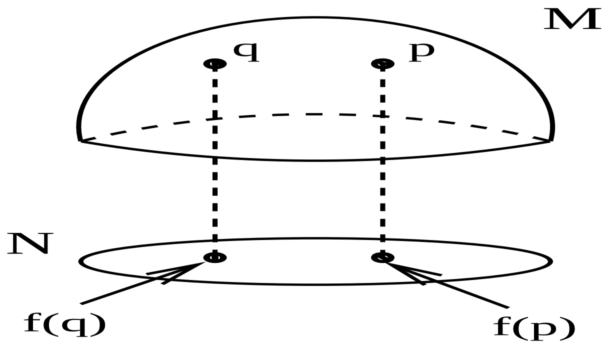
**Map into**: If a map is defined for all ponts of a manifold **M**, it is a mapping from **M into N**.

**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 $\mathbf{R}^2$, i.e. the set of all points in $\mathbf{R}^2$ 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

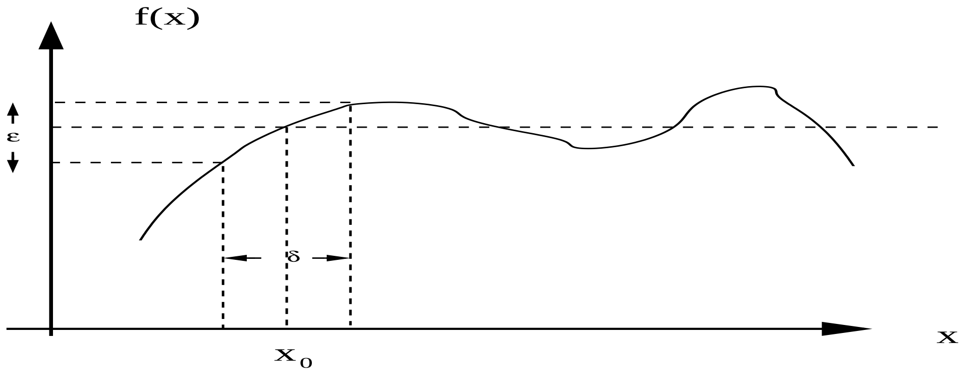
A map $f : \mathbf{M} \to \mathbf{N}$ is continuous at $x \in \mathbf{M}$ if any open set of $\mathbf{N}$ containing $f(x)$ contains the image of an open set of $\mathbf{M}$. $\mathbf{M}$ and $\mathbf{N}$ must be topological spaces, otherwise the notion of continuity has no meaning.

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 $\mathbf{R}$ to $\mathbf{R}$

$$f : \mathbf{R} \to \mathbf{R}. \tag{2.9}$$

In the elementary calculus we say that $f$ is continuous at a point $x0$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(x0)| < \epsilon, \qquad \forall x \quad \text{such that} \quad |x - x0| < \delta. \tag{2.10}$$




Let us translate this definition in terms of open sets. From the figure it is apparent that any open set containing $f(x0)$, i.e. $|f(x) - f(x0)| < r$ with $r$ arbitrary, contains an image of an open set of $\mathbf{M}$ . This is true at least in the domain of definition of $f$. This definition is more general than that of continuous functions, because we can choose $\epsilon$ and $\delta$ as big as we like, provided we remain in the domain of definition of $f$.

## 2.5 Differentiable mapping

Since we are dealing with maps of $\mathbf{R}$ into $\mathbf{R}$ (see previous figure) the notion of functions $C^k$ familiar from elementary calculus applies: if $f(x^1, x^2, ...x^n)$ is a function defined on some open region $\mathbf{S}$ of $\mathbf{R}^n$, it is said to be *differentiable of class* $C^k$ if all its partial derivatives of order less than or equal to k exist and are continuous functions on $\mathbf{S}$. For example a function $C^0$ is only continuous, and $C^\infty$ is a function such that all derivatives exist.

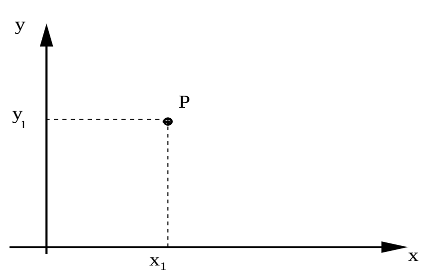
## 2.6 Manifolds

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

*A manifold is a collection of points* $\mathbf{M}$ *such that each point of* $\mathbf{M}$ *has an open neighborhood which has a continuous 1-1 map* **onto** *an open set of* $\mathbf{R}^n$*. n is the dimension of the manifold.*

---

In this definition we have used the concepts defined in the preceeding pages: the space must be topological, with no holes, i.e. 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



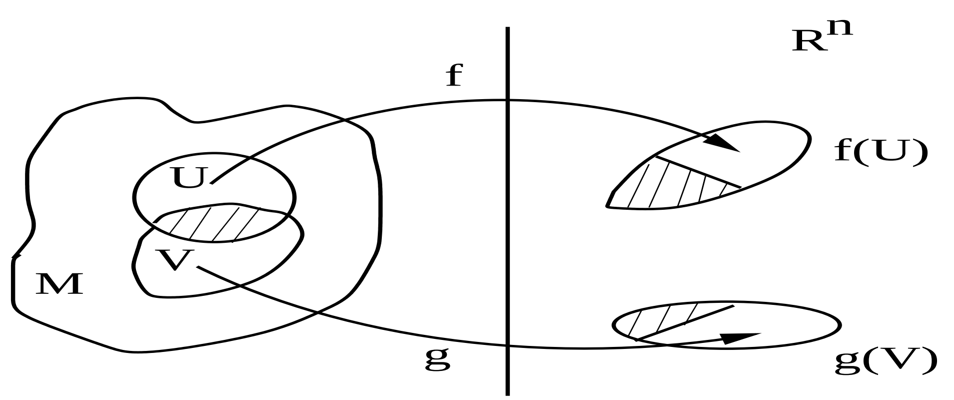

we are just using the notion of manifold: we take a point P, and map it to the point $(x^1, y^1) \in \mathbf{R}^2$ . And this operation can be done for any open neighborhood of P. It should be stressed that the definition of manifold involves open sets and not the whole of $\mathbf{M}$ and $\mathbf{R}^n$,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

DEFINITION OF COORDINATE SYSTEMS

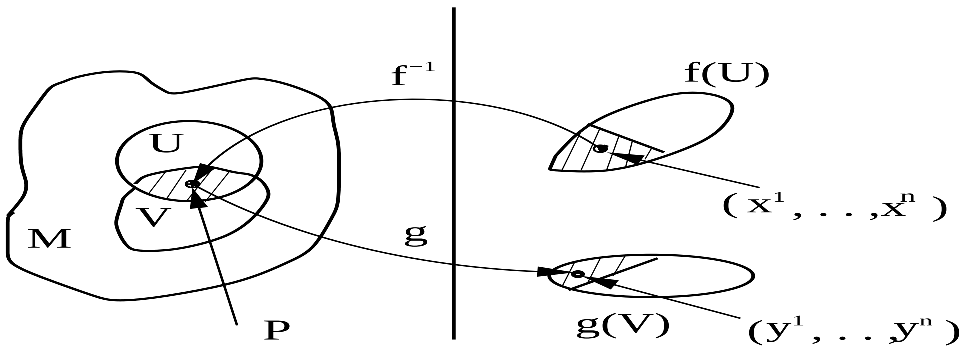
A coordinate system, or a **chart**, is a pair consisting of an open set of $\mathbf{M}$ and its map to an open set of $\mathbf{R}^n$. The open set does not necessarily include all $\mathbf{M}$ , thus there will be other open sets with the associated maps, and each point of $\mathbf{M}$ must lie in at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

AND NOW WE WANT TO MAKE A COORDINATE TRANSFORMATION.

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



The overlapping region is open (since it is the intesection 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.




Pick a point in the image of the overlapping region belonging to $f(\mathbf{U})$, say the point $(x^1, ...x^n)$. 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 $g(\mathbf{V})$, i.e. to the point $(y^1, ...y^n)$ in $\mathbf{R}^n$

$$g \circ f^{-1} : \mathbf{R}^{\backslash} \to \mathbf{R}^{\backslash}. \tag{2.11}$$

The result of this operation is a functional relation between the two sets of coordinates:

$$\begin{cases} y^1 = y^1(x^1, ...x^n) \\ . \\ . \\ y^n = y^n(x^1, ...x^n), \end{cases} \tag{2.12}$$

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 $(\mathbf{U}, f)$ and $(\mathbf{V}, g)$ 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. (2.12) can be written as

$$y^i = f^i(x^1, ...x^n), \qquad i = 1, ...n, \tag{2.13}$$

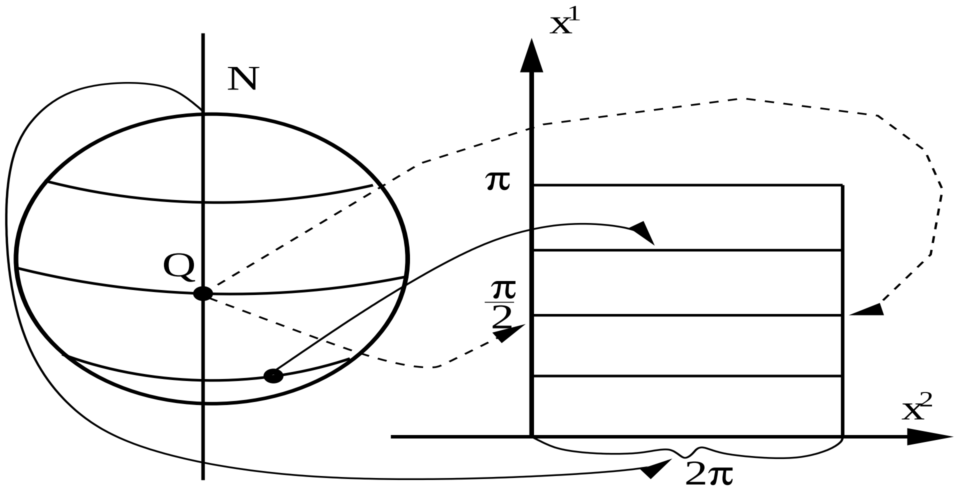
where $f^i$ are $C^k$ differentiable. Be **J** the jacobian of the transformation

$$\mathbf{J} = \frac{\partial(f^1, ... f^n)}{\partial(x^1, ... x^n)} = det \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} & . & . & . & \frac{\partial f^1}{\partial x^n} \\ \frac{\partial f^2}{\partial x^1} & \frac{\partial f^2}{\partial x^2} & . & . & . & \frac{\partial f^2}{\partial x^n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \frac{\partial f^n}{\partial x^1} & \frac{\partial f^n}{\partial x^2} & . & . & . & \frac{\partial f^n}{\partial x^n} \end{pmatrix} \tag{2.14}$$

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 $\mathbf{S}^2$). It is defined as the set of all points in $\mathbf{R}^{\ni}$ such that $(x^1)^2 + (x^2)^2 + (x^3)^2 = const.$ Suppose that we want to map the *whole* sphere to $\mathbf{R}^{\in}$ 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

$$x^1 = 0, \quad 0 \le x^2 \le 2\pi. \tag{2.15}$$

*Thus there is no map at all.*

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

$$x^2 = 0, \qquad \text{and} \qquad x^2 = 2\pi. \tag{2.16}$$

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

$$0 < x^1 < \pi, \quad 0 < x^2 < 2\pi. \tag{2.17}$$

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 lenghts.

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

In what follows we shall consider a 2-dimensional space, but the concepts we shall introduce can immediately be generalized to higher dimensional spaces, and, in particular, to a 4-dimensional spacetime. Consider for example a 2-dimensional euclidean plane. We can choose any coordinate system we like, for example cartesian coordinates $(x, y)$, polar coordinates

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan \frac{y}{x}, \end{cases} \qquad \begin{cases} x = r\cos\theta \\ y = r\sin\theta, \end{cases} \tag{3.1}$$

or arbitrary coordinates

$$\begin{cases} \xi = \xi(x, y) \qquad & d\xi = \frac{\partial\xi}{\partial x}dx + \frac{\partial\xi}{\partial y}dy \\ \eta = \eta(x, y) \qquad & d\eta = \frac{\partial\eta}{\partial x}dx + \frac{\partial\eta}{\partial y}dy. \end{cases} \tag{3.2}$$

In making a coordinate transformation it is important to impose that to any two distinct points $(x_1, y_1), (x_2, y_2)$ be assigned different pairs $(\xi_1, \eta_1), (\xi_2, \eta_2)$.

This implies that if $\Delta\xi = \Delta\eta = 0$, where $\Delta\xi$ and $\Delta\eta$ are the distances between the two points along the coordinate axes, then the two points must coincide, i.e. it must also be $\Delta x = \Delta y = 0$. This is true if the determinant of eq. (3.2), i.e. the Jacobian associated to the coordinate transformation, is non-zero

$$\mathbf{J} = \begin{pmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\xi}{\partial y} \\ \frac{\partial\eta}{\partial x} & \frac{\partial\eta}{\partial y} \end{pmatrix} \neq 0. \tag{3.3}$$

If $\mathbf{J}$ vanishes at a point, the transformation is said to be *singular* there.

## 3.1 The traditional definition of a vector

*A vector is a collection of D numbers (D = dimensions of the space) which transform like the coordinates do under a coordinate transformation.*

For example, a typical vector is the displacement

$$\vec{ds} \rightarrow_0 (dx, dy), \tag{3.4}$$

or, in a more compact form

$$\vec{ds} \rightarrow_0 \{dx^\mu\}, \quad \mu = 1, 2, \tag{3.5}$$

where the arrow indicates that $\vec{ds}$ has components with respect to a given frame $O$. If we now make a coordinate transformation, for example the transformation (3.2), the components of $\vec{ds}$ transform accordingly

$$\begin{pmatrix} d\xi \\ d\eta \end{pmatrix} = \begin{pmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\xi}{\partial y} \\ \frac{\partial\eta}{\partial x} & \frac{\partial\eta}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} dx \\ dy \end{pmatrix}. \tag{3.6}$$

If we put $(x^{1'}, x^{2'}) = (\xi, \eta)$, and $(x^1, x^2) = (x, y)$, eq. (3.6) can be rewritten in the more compact form

$$dx^{\mu'} = \sum_{\alpha=1,2} \frac{\partial x^{\mu'}}{\partial x^\alpha} dx^\alpha = \frac{\partial x^{\mu'}}{\partial x^\alpha} dx^\alpha. \tag{3.7}$$

($\alpha$ is a dummy index). Then *contravariant vectors* $\vec{A} \to_0 \{A^\mu\}$ are defined as objects that transform in the same way as $\vec{ds}$

$$A^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\beta} A^\beta, \tag{3.8}$$

where $A^{\mu'}$ are the components of the vector in the new frame. If we now define the matrix

$$(\Lambda^{\alpha'}{}_\beta) = \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{pmatrix}, \tag{3.9}$$

the transformation law can be written in the general form

$$A^{\mu'} = \Lambda^{\alpha'}{}_\beta A^\beta. \tag{3.10}$$

In addition one can define *covariant vectors* as objects that transform according to the following rule

$$A_{\mu'} = \frac{\partial x^\beta}{\partial x^{\mu'}} A_\beta = \Lambda^\beta{}_{\mu'} A_\beta, \tag{3.11}$$

where $\Lambda^\beta{}_{\mu'}$ is the inverse matrix of $\Lambda^{\beta'}{}_\mu$. However, a vector is a geometrical object. In fact it is an arrow that joins two points of a given space. We can associate components to it, and when we rotate the reference frame they 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 space)

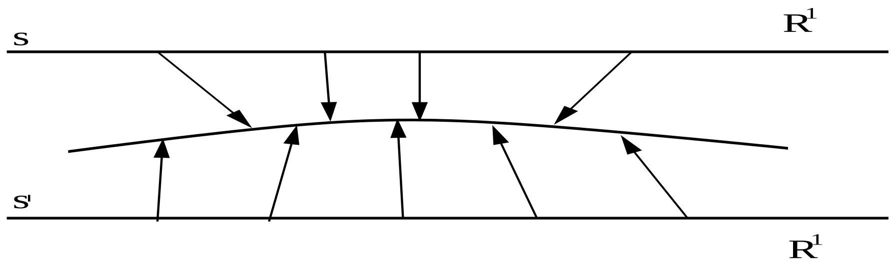
**CURVE**

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 the real line into a path in the plane. The number is called the *parameter.* For example

$$\text{curve} \quad : \{\xi = f(s), \eta = g(s), a \leq s \leq b\}, \tag{3.12}$$

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.* What happens if we change the parameter? If $s' = s'(s)$ we shall get a new curve

$$\{\xi = f'(s'), \eta = g'(s'), a' \leq s' \leq b'\}, \tag{3.13}$$




where $f', g', a', b'$ 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.

Let us now consider a curve and a differentiable function $\Phi(\xi, \eta)$. The derivative of $\Phi$ along the curve is

$$\frac{d\Phi}{ds} = \frac{\partial\Phi}{\partial\xi}\frac{d\xi}{ds} + \frac{\partial\Phi}{\partial\eta}\frac{d\eta}{ds} = \frac{\partial\Phi}{\partial x^i}\frac{dx^i}{ds}, \qquad x^1 = \xi, x^2 = \eta. \tag{3.14}$$

The set of numbers $\{\frac{dx^i}{ds}\} = (\frac{d\xi}{ds}, \frac{d\eta}{ds})$ 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

$$\vec{V} \to \{\frac{dx^i}{ds}\}, \qquad (\text{or} \qquad \vec{V} \to (\frac{d\xi}{ds}, \frac{d\eta}{ds}) \quad \text{in the two-dimensional space}) \tag{3.15}$$

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 lenght. The lenght 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.

**A vector is a geometrical object defined as the tangent vector to a given curve at a point P.** It should be remembered 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:

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 $\xi = \xi(x, y), \eta = \eta(x, y)$. The parameter $s$ is unaffected, thus

$$\begin{cases} \frac{d\xi}{ds} = \frac{d\xi}{dx}\frac{dx}{ds} + \frac{d\xi}{dy}\frac{dy}{ds} \\ \frac{d\eta}{ds} = \frac{d\eta}{dx}\frac{dx}{ds} + \frac{d\eta}{dy}\frac{dy}{ds} \end{cases} \qquad \begin{pmatrix} \frac{d\xi}{ds} \\ \frac{d\eta}{ds} \end{pmatrix} = \begin{pmatrix} \frac{d\xi}{dx} & \frac{d\xi}{dy} \\ \frac{d\eta}{dx} & \frac{d\eta}{dy} \end{pmatrix} \cdot \begin{pmatrix} \frac{dx}{ds} \\ \frac{dy}{ds} \end{pmatrix}$$

As expected, this is the same transformation as (3.10) that was used to define a contravariant vector

$$V^{\mu'} = \Lambda^{\mu'}{}_{\beta} V^{\beta}. \tag{3.16}$$

## 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 start from eq. (3.14) written for a more general n-dimensional space

$$\frac{d\Phi}{d\lambda} = \frac{\partial\Phi}{\partial x^i}\frac{dx^i}{d\lambda}, \qquad\qquad i = 1, ....n. \tag{3.17}$$

This is the directional derivative along a curve parametrized with the parameter $\lambda$, and it has been used to define vectors. Since the function $\Phi$ is totally arbitrary, we can rewrite this expression as

$$\frac{d}{d\lambda} = \frac{dx^i}{d\lambda}\frac{\partial}{\partial x^i}, \tag{3.18}$$

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

$$\frac{d}{d\lambda} = \frac{dx^i}{d\lambda}\frac{\partial}{\partial x^i}, \qquad \frac{d}{d\mu} = \frac{dx^i}{d\mu}\frac{\partial}{\partial x^i}. \tag{3.19}$$

$\{\frac{dx^i}{d\mu}\}$ are the components of the vector tangent to the second curve. Take the linear combination

$$a\frac{d}{d\lambda} + b\frac{d}{d\mu} = (a\frac{dx^i}{d\lambda} + b\frac{dx^i}{d\mu})\frac{\partial}{\partial x^i}. \tag{3.20}$$

The numbers $(a\frac{dx^i}{d\lambda} + b\frac{dx^i}{d\mu})$ 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

$$\frac{d}{ds} = (a\frac{dx^i}{d\lambda} + b\frac{dx^i}{d\mu})\frac{\partial}{\partial x^i}, \tag{3.21}$$

and consequently at P it must be

$$a) \qquad \frac{d}{ds} = a\frac{d}{d\lambda} + b\frac{d}{d\mu}. \tag{3.22}$$

In addition it is easy to verify that the directional derivative satisfies the following properties

$$1) \quad a(\frac{d}{d\lambda} + \frac{d}{d\mu}) = a\frac{d}{d\lambda} + a\frac{d}{d\mu} \tag{3.23}$$

$$2) \quad (a+b)\frac{d}{d\lambda} = a\frac{d}{d\lambda} + b\frac{d}{d\lambda} \tag{3.24}$$

$$3) \quad (ab)\frac{d}{d\lambda} = a(b\frac{d}{d\lambda}) \tag{3.25}$$

$$4) \quad 1\cdot\frac{d}{d\lambda} = \frac{d}{d\lambda} \tag{3.26}$$

Thus the set of directional derivatives is a vector space because it satisfies (1-4) and (a).

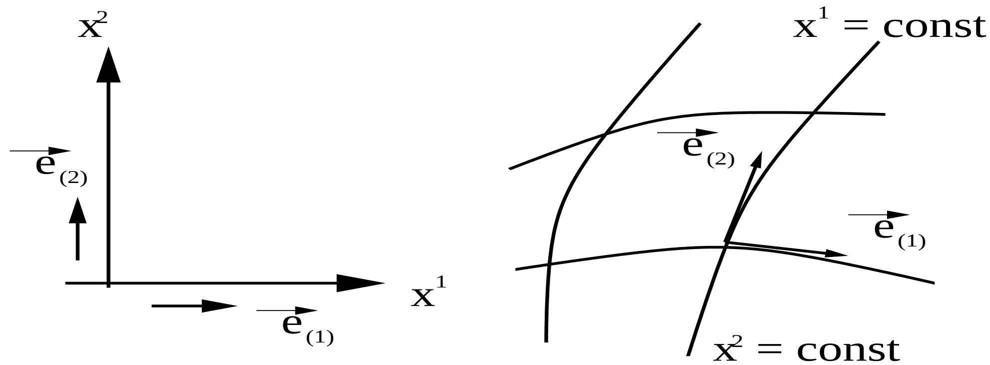
In any coordinate system there are special curves, the coordinates lines (think for example to the grid of cartesian coordinates). The derivatives along these lines are simply $\frac{\partial}{\partial x^i}$. Eq. (3.18) shows that the generic directional derivative $\frac{d}{d\lambda}$ can always be expressed as a linear combination of $\frac{\partial}{\partial x^i}$. It follows that $\frac{\partial}{\partial x^i}$ are a basis for this vector space. Then eq. (3.18) shows that $\{\frac{dx^i}{d\lambda}\}$ 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 matematicians often say that $\frac{d}{d\lambda}$ is the vector tangent to the curve $x^i(\lambda)$.

*Vectors do not lie in* $\mathbf{M}$, *but in the tangent space to M, called* $\mathbf{T_P}$ 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 $\mathbf{T_P}$ . In any event $\mathbf{T_P}$ *has the same dimensions as the manifold* $\mathbf{M}$.

Any collection of n linearly independent vectors of $\mathbf{T_P}$ is a basis for $\mathbf{T_P}$ . The most natural choice is the following: given a coordinate system $\{x^i\}$ in a neighborhood U of a point P, the coordinates define a basis $\left\{\frac{\partial}{\partial x^i}\right\}$, called *the coordinate basis*, and any vector $\vec{V}$ at the point P, can be expressed as

$$\vec{V} = \sum_i V^i \frac{\partial}{\partial x^i}. \tag{3.27}$$

In other words, the basis vectors are tangent to the coordinate lines.




If we choose a different basis $\{\vec{e}_{j'}\}$

$$\vec{V} = \sum_{j'} V^{j'} \vec{e}_{j'}. \tag{3.28}$$

The numbers $\{V^i\}$ are the components of $\vec{V}$ with respect to the basis $\left\{\frac{\partial}{\partial x^i}\right\}$, while $\{V^{j'}\}$ are the components of $\vec{V}$ on $\{\vec{e}_j\}$. When we write $\vec{e}_{j'}$, the $j'$ does not indicate a component of the basis vector $\vec{e}_{j'}$, but which vector of the basis we choose. It is better to indicate this index in parenthesis $e_{(j')}$ i.e.

$$\vec{V} = \sum_{j'} V^{j'} \vec{e}_{(j')}. \tag{3.29}$$

EXAMPLE

Consider the 4-dimensional flat spacetime of Special Relativity, and let us restrict to the (x-y) plane, where we choose the coordinates

$$(ct, x, y) \equiv (x^0, x^1, x^2) \tag{3.30}$$

The coordinate basis is the set of vectors

$$\begin{aligned} \frac{\vec{\partial}}{\partial x^0} &= \vec{e}_{(0)} \to \quad (1,0,0) \\ \frac{\vec{\partial}}{\partial x^1} &= \vec{e}_{(1)} \to \quad (0,1,0) \\ \frac{\vec{\partial}}{\partial x^2} &= \vec{e}_{(2)} \to \quad (0,0,1), \end{aligned} \tag{3.31}$$

or, in a compact form

$$e^{\beta}_{(\alpha)} = \delta^{\beta}_{\alpha}. \tag{3.32}$$

(The superscript $\beta$ now indicates the $\beta$-component of the $\alpha$-th vector). In this basis any vector $\vec{A}$ can be written as

$$\vec{A} = A^0\vec{e}_{(0)} + A^1\vec{e}_{(1)} + A^2\vec{e}_{(2)} = A^{\alpha}\vec{e}_{(\alpha)}. \tag{3.33}$$

$\{A^i\} = (A^0, A^1, A^2)$ are the components of $\vec{A}$ with respect to the coordinate basis. Let us put

$$\vec{e}_{(0)} = \vec{e}_t, \qquad \vec{e}_{(1)} = \vec{e}_x, \qquad \vec{e}_{(2)} = \vec{e}_y. \tag{3.34}$$

Let us consider the following coordinate transformation

$$\begin{cases} (x^0, x, y) \to (x^0, r, \theta) \\ x^0 = x^{0'} \\ x = r\cos\theta \\ y = r\sin\theta, \end{cases} \tag{3.35}$$

A new coordinate basis will be associated to the new coordinates

$$\frac{\vec{\partial}}{\partial x^{0'}} = \vec{e}_{(0')} = \vec{e}_{(0)} \tag{3.36}$$
$$\frac{\vec{\partial}}{\partial x^{1'}} = \vec{e}_{(1')} = \vec{e}_r \tag{3.37}$$
$$\frac{\vec{\partial}}{\partial x^{2'}} = \vec{e}_{(2')} = \vec{e}_\theta.$$

Each of these vectors will be a linear combination of the old vectors $\vec{e}_t, \vec{e}_x, \vec{e}_y$. To find a general rule we observe that a vector $\vec{V}$ can be written as

$$\vec{V} = V^j\vec{e}_{(j)}, \qquad \text{in terms of the old basis} \tag{3.38}$$
$$\vec{V} = V^{j'}\vec{e}_{(j')}, \qquad \text{in terms of the new basis,}$$

and consequently

$$V^i\vec{e}_{(i)} = V^{i'}\vec{e}_{(i')}. \tag{3.39}$$

$V^j$ $V^{j'}$ are the components of $\vec{V}$ respectively in the old and in the new basis. If we now substitute eq. (3.16) which we report here again

$$V^{i'} = \Lambda^{i'}{}_k V^k \tag{3.40}$$

into eq. (3.39) we find

$$V^i \vec{e}_{(i)} = \Lambda^{i'}{}_k V^k \vec{e}_{(i')}. \tag{3.41}$$

By relabelling the dummy indices this equation can be written as

$$\left[\Lambda^{i'}{}_k \vec{e}_{(i')} - \vec{e}_{(k)}\right] V^k = 0, \tag{3.42}$$

i.e.

$$\vec{e}_{(k)} = \Lambda^{i'}{}_k \vec{e}_{(i')}. \tag{3.43}$$

Multiplying both members by $\Lambda^j{}_{i'}$ and remembering that

$$\Lambda^j{}_{i'} \Lambda^{i'}{}_k = \frac{\partial x^j}{\partial x^{i'}} \frac{\partial x^{i'}}{\partial x^k} = \frac{\partial x^j}{\partial x^k} = \delta^j_k, \tag{3.44}$$

$$\vec{e}_{(i')} = \Lambda^k{}_{i'} \vec{e}_{(k)}. \tag{3.45}$$

This is the transformation law we were looking for. Summarizing:

$$\begin{cases} \vec{e}_{(k)} = \Lambda^{i'}{}_k \vec{e}_{(i')}, \\ \vec{e}_{(i')} = \Lambda^k{}_{i'} \vec{e}_{(k)}. \end{cases} \tag{3.46}$$

We are now in a position to compute the new basis vectors in terms of the old ones:

$$\vec{e}_{(0')} = \Lambda^k{}_{0'} \vec{e}_{(k)}, \qquad \Lambda^k{}_{0'} = \frac{\partial x^k}{\partial x^{0'}}. \tag{3.47}$$

In the example we are considering only $\Lambda^0{}_{0'} \neq 0$ and it is equal to 1. It follows that

$$\vec{e}_{(0)} = \vec{e}_{(0')}. \tag{3.48}$$

In addition

$$\vec{e}_{(1')} = \vec{e}_r = \Lambda^k{}_r \vec{e}_{(k)}, \tag{3.49}$$

and since

$$\begin{aligned} \Lambda^0{}_r &= \frac{\partial x^0}{\partial r} = 0, \\ \Lambda^1{}_r &= \frac{\partial x^1}{\partial r} = \frac{\partial x}{\partial r} = \cos\theta, \\ \Lambda^2{}_r &= \frac{\partial x^2}{\partial r} = \frac{\partial y}{\partial r} = \sin\theta, \end{aligned} \tag{3.50}$$

$$\vec{e}_r = \cos\theta \vec{e}_x + \sin\theta \vec{e}_y. \tag{3.51}$$

Similarly

$$\vec{e}_{(2')} = \vec{e}_\theta = \Lambda^k{}_\theta \vec{e}_{(k)}, \tag{3.52}$$

and since

$$\begin{aligned} \Lambda^0{}_\theta &= 0, \\ \Lambda^1{}_\theta &= \frac{\partial x^1}{\partial \theta} = \frac{\partial x}{\partial \theta} = -r\sin\theta, \\ \Lambda^2{}_\theta &= \frac{\partial x^2}{\partial \theta} = \frac{\partial y}{\partial \theta} = r\cos\theta, \end{aligned} \tag{3.53}$$

hence

$$\vec{e}_\theta = -r\sin\theta \vec{e}_x + r\cos\theta \vec{e}_y. \tag{3.54}$$

## 3.4 One-forms

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

Properties of one-forms.

1) Linearity

$$\tilde{\omega}(a\vec{V} + b\vec{W}) = a\tilde{\omega}(\vec{V}) + b\tilde{\omega}(\vec{W}). \tag{3.55}$$

2) Multiplication by real numbers

$$(a\tilde{\omega})(\vec{V}) = a[\tilde{\omega}(\vec{V})] \tag{3.56}$$

3) Addition

$$[\tilde{\omega} + \tilde{\sigma}](\vec{V}) = \tilde{\omega}(\vec{V}) + \tilde{\sigma}(\vec{V}). \tag{3.57}$$

Since one-forms satisfy the axioms (3.55-3.57), they form a vector space, which is called the **dual** vector space to $\mathbf{T_P}$ , and it is indicated as $\mathbf{T_P^*}$ . The reason why $\mathbf{T_P^*}$ is called dual to $\mathbf{T_P}$ is that also vectors can be regarded as linear, real valued functions of one-forms: a vector $\vec{V}$ takes a 1-form $\tilde{\omega}$ and associates a number to it, which we call $\vec{V}(\tilde{\omega})$. If we rewrite eq. (3.55) we see that

$$(a\tilde{\omega} + b\tilde{\sigma})(\vec{V}) = (a\tilde{\omega})(\vec{V}) + (b\tilde{\sigma})(\vec{V}) = a\tilde{\omega}(\vec{V}) + b\tilde{\sigma}(\vec{V}) = \tag{3.58}$$

$$a\left[\text{value of} \quad \vec{V} \quad \text{on} \quad \tilde{\omega}\right] + b\left[\text{value of} \quad \vec{V} \quad \text{on} \quad \tilde{\sigma}\right] =$$

$a\vec{V}(\tilde{\omega}) + b\vec{V}(\tilde{\sigma}).$

Thus it is the linearity which allows the "duality": each (vectors or 1-forms) can be regarded as functions of the other (1-forms or vectors) to produce a number. We can equivalently say

$$\tilde{\omega}(\vec{V}) \equiv \vec{V}(\tilde{\omega}) \equiv < \tilde{\omega}, \vec{V} >, \tag{3.59}$$

in the sense that the three "operations" give as a result the same number. This point will be further clarified in the following.

Since $\mathbf{T}^*_{\mathbf{p}}$ is dual to $\mathbf{T}_{\mathbf{p}}$ , once we choose a basis for vectors, say $\{\vec{e}_{(i)}, i = 1, n\}$, we can introduce a dual basis for one-forms defined as follows: *the "preferred" dual basis* $\{\tilde{\omega}^{(i)}, i = 1, n\}$, takes any vector $\vec{V}$ in $\mathbf{T}_{\mathbf{p}}$ and produces its components

$$\tilde{\omega}^{(i)}(\vec{V}) = V^i. \tag{3.60}$$

It should be remembered that an index in parenthesis does not refer to a component, but selects the -th vector, or one-form, 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

$$\tilde{\omega}^{(i)}(\vec{V} + \vec{W}) = V^i + W^i, \tag{3.61}$$

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 one of the basis vectors $\vec{e}_{(j)}$, since only the j-th component of $\vec{e}_{(j)}$ is different from zero and equal to 1 (if it is a coordinate basis), we have

$$\tilde{\omega}^{(i)}(\vec{e}_{(j)}) = \delta^i_j. \tag{3.62}$$

We now want to answer the questions:

1) who tells us that $\{\tilde{\omega}^{(i)}\}$ are linearly independent?

2) can we define the components of a 1-form as we define the components of a vector?

Consider any one-form $\tilde{q}$ acting on an arbitrary vector $\vec{V}$

$$\begin{aligned} \tilde{q}(\vec{V}) = & \quad \tilde{q}(\sum_j V^j \vec{e}_{(j)}) = \sum_j V^j \tilde{q}(\vec{e}_{(j)}) = \\ = & \quad \sum_j \tilde{\omega}^{(j)}(\vec{V}) \tilde{q}(\vec{e}_{(j)}), \end{aligned} \tag{3.63}$$

where the third equality follows from the fact that $V^j$ are numbers, and the last from eq.(3.60). Thus $\tilde{q}(\vec{V})$ is expressed as a linear combination of basis one-forms $\tilde{\omega}^{(j)}$ applied to $\vec{V}$, and the numbers

$$q_j = \tilde{q}(\vec{e}_{(j)}) \tag{3.64}$$

are defined to be *the components of* $\tilde{q}$ *on the basis* $\{\tilde{\omega}^{(i)}\}$ *dual to* $\vec{e}_{(j)}$. Consequently we can write

$$\tilde{q}(\vec{V}) = \sum_j q_j \tilde{\omega}^{(j)}(\vec{V}). \tag{3.65}$$

Since $\vec{V}$ is arbitrary, we can also write

$$\tilde{q} = \sum_j q_j \tilde{\omega}^{(j)}. \tag{3.66}$$

Since any one-form can be expressed as a linear combination of the n basis one-forms $\tilde{\omega}^{(j)}$ they form a basis.

Consider a region of space U and choose a coordinate system $\{x^i\}$ We have seen that this defines a *natural coordinate basis for vectors* $\{\frac{\partial}{\partial x^{(i)}}\}$ For what we said before, it also defines a *natural set of coordinate basis one-forms*, indicated as $\{\tilde{d}x^{(i)}\}$ whose components are

$$(\tilde{d}x^{(i)})_j = \tilde{d}(\frac{\partial}{\partial x^{(j)}}) \equiv \frac{\partial x^{(i)}}{\partial x^{(j)}} = \delta^i_j. \tag{3.67}$$

And now the most important thing. From eq. (3.65), since $\tilde{\omega}^{(j)}(\vec{V}) = V^j$, we find

$$\tilde{q}(\vec{V}) = \sum_j q_j V^j, \qquad \text{(this operation is called contraction)} \tag{3.68}$$

which tells us, if we know the components of $\tilde{q}$ and $\vec{V}$ , how to compute the number which results from the application of $\tilde{q}$ on $\vec{V}$ .

---

From eq. (3.68) 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 $\vec{V}$

$$\tilde{q}(\vec{V}) = q_j V^j = q_0 V^0 + q_1 V^1 + ... q_n V^n = \quad \text{a number.} \tag{3.69}$$

Due to the linearity of the previous equation, we can also say that

$$q_j V^j = q_0 V^0 + q_1 V^1 + ... q_n V^n = \vec{V}(\tilde{q}), \tag{3.70}$$

and define vectors as those linear functions that, when applied to one-forms, produce a number.

---

Let us now make a coordinate transformation $x^{\mu'} = x^{\mu'}(x^i)$. We can ask two things

1) How do the components $q_j$ of a one-form $\tilde{q}$ change?

2) Will the new basis 1-forms be a linear combination of the old ones, and which combination?

1) By definition

$$q_j = \tilde{q}(\vec{e}_{(j)}). \tag{3.71}$$

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

$$\vec{e}_{(i')} = \Lambda^k{}_{i'}\vec{e}_{(k)}, \tag{3.72}$$

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

$$q_{j'} = \tilde{q}(\vec{e}_{(j')}) = \tilde{q}[\Lambda^k{}_{j'}\vec{e}_{(k)}] = \Lambda^k{}_{j'}\tilde{q}(\vec{e}_{(k)}) = \Lambda^k{}_{j'}q_k, \tag{3.73}$$

hence

$$q_{j'} = \Lambda^k{}_{j'}q_k. \tag{3.74}$$

If we compare this result with eq. (3.11) we immediately recognize that this is the way covariant vectors transform, thus *covariant vectors are one-forms.*

2) We now want to see, as we did in the case of vectors, whether the new basis one-forms can be expressed as a linear combination of the old ones. We shall proceed along the same lines. From eq. (3.66) we see that

$$\tilde{q} = q_j \tilde{\omega}^{(j)} = q_{k'} \tilde{\omega}^{(k')}, \tag{3.75}$$

(sum removed according to Einstein's convenction), where $\{\tilde{\omega}^{(k')}\}$ are the new basis one-forms. But

$$q_{k'} = \Lambda^i{}_{k'} q_i, \tag{3.76}$$

therefore

$$q_j \tilde{\omega}^{(j)} = \Lambda^i{}_{k'} q_i \tilde{\omega}^{(k')}. \tag{3.77}$$

This equation can be rewritten as

$$[\Lambda^i{}_{k'} \tilde{\omega}^{(k')} - \tilde{\omega}^{(i)}] q_i = 0, \tag{3.78}$$

hence

$$\tilde{\omega}^{(i)} = \Lambda^i{}_{k'} \tilde{\omega}^{(k')}. \tag{3.79}$$

The matrix $\Lambda^i{}_{j'}$ is inverse of $\Lambda^{k'}{}_i$. Thus

$$\Lambda^{k'}{}_j \Lambda^j{}_{i'} = \delta^{k'}_{i'}, \qquad \text{or} \qquad \Lambda^{k'}{}_j \Lambda^i{}_{k'} = \delta^i_j. \tag{3.80}$$

Multiplying both sides of eq. (3.79) by $\Lambda^{k'}{}_j$ we find

$$\Lambda^{k'}{}_j \tilde{\omega}^{(i)} = \Lambda^{k'}{}_j \Lambda^i{}_{k'} \tilde{\omega}^{(k')} = \delta^i_j \tilde{\omega}^{(k')}, \tag{3.81}$$

hence

$$\tilde{\omega}^{(k')} = \Lambda^{k'}{}_j \tilde{\omega}^{(j)}. \tag{3.82}$$

Summarizing

$$\begin{cases} \tilde{\omega}^{(i)} = \Lambda^{i}{}_{k'}\tilde{\omega}^{(k')} \\ \tilde{\omega}^{(k')} = \Lambda^{k'}{}_{j}\tilde{\omega}^{(j)} \end{cases} \tag{3.83}$$

EXAMPLE

Let us consider the same coordinate transformation analyzed in sec. 3.4. We start with minkowskian coordinates $(x^0, x, y)$. The coordinate basis for vectors is $\{\frac{\partial}{\partial x^{(i)}}\}$ and the dual basis for one-forms is $\{\tilde{d}x^{(i)}\}$

$$\tilde{d}x^{(0)} \to (1, 0, 0) \equiv \tilde{d}t \tag{3.84}$$

$$\tilde{d}x^{(1)} \to (0, 1, 0) \equiv \tilde{d}x \tag{3.85}$$

$$\tilde{d}x^{(2)} \to (0, 0, 1) \equiv \tilde{d}y \tag{3.86}$$

If we now change to polar coordinates, according to eq. (3.82) we find

$$\tilde{\omega}^{(0')} = \tilde{d}t' = \Lambda^{0'}{}_{j}\tilde{d}x^{(j)}. \tag{3.87}$$

Since $\Lambda^{0'}{}_{j} = \frac{\partial x^{0'}}{\partial x^{j}}$, only $\Lambda^{0'}{}_{0} \neq 0$, and equal to unity, thus

$$\tilde{\omega}^{(0')} = \tilde{d}t' = \tilde{d}x^{(0)}. \tag{3.88}$$

Similarly

$$\tilde{\omega}^{(1')} = \Lambda^{1'}{}_{j}\tilde{d}x^{(j)} = \frac{\partial x^{1'}}{\partial x^{j}}\tilde{d}x^{(j)} = \frac{\partial r}{\partial x}\tilde{d}x^{(1)} + \frac{\partial r}{\partial y}\tilde{d}x^{(2)}. \tag{3.89}$$

Since

$$\frac{\partial r}{\partial x} = \frac{x}{r} = \cos\theta, \qquad \text{and} \qquad \frac{\partial r}{\partial y} = \frac{y}{r} = \sin\theta \tag{3.90}$$

it follows that

$$\tilde{d}r = \cos\theta \tilde{d}x + \sin\theta \tilde{d}y. \tag{3.91}$$

Moreover

$$\tilde{\omega}^{(2')} = \tilde{d}\theta = \frac{\partial\theta}{\partial x}\tilde{d}x^{(1)} + \frac{\partial\theta}{\partial y}\tilde{d}x^{(2)}, \tag{3.92}$$

hence

$$\tilde{d}\theta = -\frac{1}{r}\sin\theta \tilde{d}x + \frac{1}{r}\cos\theta \tilde{d}y. \tag{3.93}$$

---

AN EXAMPLE OF ONE-FORM.

Consider a scalar field $\Phi(x^0, x^1, ...x^n)$. The gradient of a scalar field is

$$\tilde{\Phi} \to (\frac{\partial\Phi}{\partial x^0}, \frac{\partial\Phi}{\partial x^1}, ..., \frac{\partial\Phi}{\partial x^n}). \tag{3.94}$$

It is easy to see, for example, that the components transform according to eq. (3.74), in fact

$$\tilde{\Phi}_j = \frac{\partial\Phi}{\partial x^j}, \qquad \text{and} \qquad \tilde{\Phi}_{j'} = \frac{\partial\Phi}{\partial x^{j'}} = \frac{\partial\Phi}{\partial x^k} \cdot \frac{\partial x^k}{\partial x^{j'}}; \tag{3.95}$$

since $\Lambda^k{}_{j'} = \frac{\partial x^k}{\partial x^{j'}}$, it follows that

$$\tilde{\Phi}_{j'} = \Lambda^k{}_{j'}\tilde{\Phi}_k, \tag{3.96}$$

same as eq. (3.74). Thus the gradient of a scalar field is a one-form.

# 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 a manifold $M$. A tensor of type $\begin{pmatrix} N \\ N' \end{pmatrix}$ at $P$ is defined to be *a linear function which takes as arguments $N$ one-forms and $N'$ vectors and associates a number to them.*

For example if $F$ is a $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ tensor this means that

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

is a number and the linearity implies that

$$F(a\tilde{\omega} + b\tilde{g}, \tilde{\sigma}, \vec{V}, \vec{W}) = aF(\tilde{\omega}, \tilde{\sigma}, \vec{V}, \vec{W}) + bF(\tilde{g}, \tilde{\sigma}, \vec{V}, \vec{W})$$

and

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

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

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

In the same way

$$F(\tilde{\omega},\tilde{g},\vec{V}\vec{W}) \neq F(\tilde{g},\tilde{\omega},\vec{V},\vec{W}). \tag{4.2}$$

EXAMPLES

A $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 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 ia a one-form.

$$\tilde{q}(\vec{V}) = \sum_j q_j V^j \equiv q_j V^j. \tag{4.3}$$

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

$$\vec{V}(\tilde{q}) = q_j V^j. \tag{4.4}$$

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.

We know how to associate to each vector a number: we contract it with a one-form according to 4.3. The obvious generalization in the case of two vectors is

$$F(\vec{V}, \vec{W}) = \tilde{p}(\vec{V})\tilde{q}(\vec{W}) = p_j V^j q_k W^k.$$

where $\tilde{p}$ and $\tilde{q}$ are two arbitrary one-forms, and $\vec{V}$ and $\vec{W}$ are two arbitrary vectors. This operation is indicated with the following symbol $\tilde{p} \otimes \tilde{q}$, where $\otimes$ is called the "outer product sign" and means precisely that if $\tilde{p}$ and $\tilde{q}$ have arguments $\vec{V}$ and $\vec{W}$ the result is the number $p_j q_k V^j W^k$.

How do we compute the components of a $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor? Let us recall what we did with one-forms (see eqs. 3.64 and 3.65).

The components of a one-form $\tilde{q}$ are defined as the numbers that we obtain when the argument of the one-form is one of the basis vectors

$$q_j = \tilde{q}(\vec{e}_{(j)}).$$

It is now very easy to generalize these definitions to the case of a $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor: the components of a $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor will be the numbers that are produced when the arguments are the basis vectors

$$F_{\alpha\beta} = F(\vec{e}_{(\alpha)}, \vec{e}_{(\beta)}). \tag{4.5}$$

Since $\vec{e}_{(\alpha)}$ and $\vec{e}_{(\beta)}$ have $n$ components each, $F_{\alpha\beta}$ will be an $n \times n$ matrix. So now if we take as an arguments of $F$ two arbitrary vectors we find

$$\begin{aligned} F(\vec{A}, \vec{B}) &= F(A^\alpha \vec{e}_{(\alpha)}, B^\beta \vec{e}_{(\beta)}) = \\ &= A^\alpha B^\beta F(\vec{e}_{(\alpha)}, \vec{e}_{(\beta)}) = \\ &= F_{\alpha\beta} A^\alpha B^\beta. \end{aligned} \tag{4.6}$$

It is now clear what does it mean to say that $F$ is a linear function that associates a number to two vectors: the number is $F_{\alpha\beta} A^\alpha A^\beta$.

Can we construct a basis for $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensors as we did for one-forms?

We want to write

$$F = F_{\alpha\beta} \tilde{\omega}^{(\alpha)(\beta)} \tag{4.7}$$

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

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

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

On the other hand, we know that $A^\alpha = \tilde{\omega}^{(\alpha)}(\vec{A})$ and $B^\beta = \tilde{\omega}^{(\beta)}(\vec{B})$ and therefore eq. (4.6) can be written as

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

By equating eqs. (4.8) and (4.9) we find

$$\tilde{\omega}^{(\alpha)(\beta)}(\vec{A}, \vec{B}) = \tilde{\omega}^{(\alpha)}(\vec{A})\tilde{\omega}^{(\beta)}(\vec{B}),$$

or, by the definition of outer product

$$\tilde{\omega}^{(\alpha)(\beta)} = \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)}. \tag{4.10}$$

Thus the basis for $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensors can be constructed by taking the outer product of the basis one-forms. Finally

$$F = F_{\alpha\beta}\tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)} \tag{4.11}$$

## 4.2 Symmetries

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

$$F(\vec{A}, \vec{B}) = F(\vec{B}, \vec{A}) \quad \forall \vec{A}, \vec{B}. \tag{4.12}$$

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

$$F_{\alpha\beta}A^{\alpha}A^{\beta} = F_{\beta\alpha}A^{\beta}A^{\alpha}, \Rightarrow F_{\alpha\beta} = F_{\beta\alpha} \tag{4.13}$$

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

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

$$F_{(s)}(\vec{A}, \vec{B}) = \frac{1}{2}[F(\vec{A}, \vec{B}) + F(\vec{B}, \vec{A})]. \tag{4.14}$$

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

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

and consequently the components of the symmetric tensor are

$$F_{(s)\alpha\beta} = \frac{1}{2}[F_{\alpha\beta} + F_{\beta\alpha}]. \tag{4.15}$$

The components of a symmetric tensor are often indicated as

$$F_{(\alpha\beta)} = \frac{1}{2}[F_{\alpha\beta} + F_{\beta\alpha}]. \tag{4.16}$$

A $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor is **antisymmetric** if

$$F(\vec{A}, \vec{B}) = -F(\vec{B}, \vec{A}) \qquad \forall \vec{A}, \vec{B}, \qquad \text{i.e.} \qquad F_{\alpha\beta} = -F_{\beta\alpha}. \tag{4.17}$$

Again from any $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ 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

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

also indicated as

$$F_{[\alpha\beta]} = \frac{1}{2}[F_{\alpha\beta} - F_{\beta\alpha}]. \tag{4.18}$$

It is clear that any tensor $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ can be written as the sum of its symmetric and antisymmetric part

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

It is now clear that we can construct any sort of tensor using the procedure that we have developed in the previous pages. Thus for example a $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ tensor $T(\tilde{\alpha}, \tilde{\sigma})$ is a function that associates to two one-forms $\tilde{\alpha}$ and $\tilde{\sigma}$ a number. Since we know that if we contract a vector with a one-form we get a number the tensor $T$ must be given by the outer product of two vectors

$$T(\tilde{\alpha}, \tilde{\sigma}) = \vec{V}(\tilde{\alpha})\vec{W}(\tilde{\sigma}) = V^j \alpha_j W^k \sigma_k,$$

or

$$T = \vec{V} \otimes \vec{W}. \tag{4.19}$$

But now (compare with eq. 4.6)

$$T(\tilde{\alpha}, \tilde{\sigma}) = T(\alpha_j \tilde{\omega}^{(j)}, \sigma_k \tilde{\omega}^{(k)}) = \alpha_j \sigma_k T(\tilde{\omega}^{(j)}, \tilde{\omega}^{(k)}) = T^{jk} \alpha_j \sigma_k, \tag{4.20}$$

where $T^{jk}$ are the components of $T$

$$T^{jk} = T(\tilde{\omega}^{(j)}, \tilde{\omega}^{(k)}), \tag{4.21}$$

and the basis for this tensor will be

$$T = T^{\alpha\beta} \vec{e}_\alpha \otimes \vec{e}_\beta. \tag{4.22}$$

(It can be found by using the procedure that led to eq.4.11).

---

Exercise: prove that the $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor $\vec{V} \otimes \tilde{\sigma}$ has components $V^j \sigma_j$ and find the basis for $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 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 $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor

$$F = F_{\alpha\beta} \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)} \tag{4.23}$$

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

$$\tilde{\omega}^{(i)} = \Lambda^{i}{}_{k'}\tilde{\omega}^{(k')} \quad , \quad \tilde{\omega}^{(i')} = \Lambda^{i'}{}_{k}\tilde{\omega}^{k} \tag{4.24}$$

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

$$F = F_{\alpha'\beta'}\tilde{\omega}^{(\alpha')} \otimes \tilde{\omega}^{(\beta')}. \tag{4.25}$$

By equating (4.23) and (4.25)

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

Replacing $\tilde{\omega}^{(\alpha)}$ and $\tilde{\omega}^{(\beta)}$ by using the first of eqs. 4.24

$$F_{\alpha'\beta'}\tilde{\omega}^{(\alpha')} \otimes \tilde{\omega}^{(\beta')} = F_{\alpha\beta}\Lambda^{\alpha}{}_{k'}\tilde{\omega}^{(k')} \otimes \Lambda^{\beta}{}_{i'}\tilde{\omega}^{(i')} = F_{\alpha\beta}\Lambda^{\alpha}{}_{k'}\Lambda^{\beta}{}_{i'}\tilde{\omega}^{(k')} \otimes \tilde{\omega}^{(i')},$$

or by relabelling the dummy indices

$$F_{k'i'}\tilde{\omega}^{(k')} \otimes \tilde{\omega}^{(i')} = F_{\alpha\beta}\Lambda^{\alpha}{}_{k'}\Lambda^{\beta}{}_{i'}\tilde{\omega}^{(k')} \otimes \tilde{\omega}^{(i')},$$

and finally

$$F_{k'i'} = F_{\alpha\beta}\Lambda^{\alpha}{}_{k'}\Lambda^{\beta}{}_{i'}, \tag{4.26}$$

or, by writing explicitely the elements of the matrix $\Lambda^{\alpha}{}_{k'}$

$$F_{k'i'} = F_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial x^{k'}}\frac{\partial x^{\beta}}{\partial x^{i'}}, \tag{4.27}$$

where $\{x^{i'}\}$ are the new coordinates.

In a similar way, by using eqs. 3.43 and 3.45 we would find that

$$T^{k'i'} = T^{\alpha\beta}\Lambda^{k'}{}_{\alpha}\Lambda^{i'}{}_{\beta}, \tag{4.28}$$

and

$$T^{k'}{}_{i'} = T^{\alpha}{}_{\beta}\Lambda^{k'}{}_{\alpha}\Lambda^{\beta}{}_{i'} \tag{4.29}$$

IMPORTANT

I would like to stress the following point: the notion of a tensor that we have introduced is independent on 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 (4.23) and (4.25). The operations that we are allowed to make with tensors are the following. Given a tensor **T** and its components $\{T^{i..}_{j..}\}$ on some basis, we can multiply each component by a number $a$ and find $\{aT^{i..}_{j..}\}$. These are the components of the tensor **aT** of the same type.

Further operations are

1) Addition and subtraction of tensors of the same type

$$T + G = W \qquad \rightarrow \qquad T_{\alpha\beta} + G_{\alpha\beta} = W_{\alpha\beta}$$

2) If we multiply the components of two tensors, we find a tensor whose

type is the sum of the two

$$\begin{array}{ccccc} G^{\alpha\beta} & \cdot & T_{k\delta} & = & W^{\alpha\beta}{}_{k\delta} \\ \begin{pmatrix} 2 \\ 0 \end{pmatrix} & & \begin{pmatrix} 0 \\ 2 \end{pmatrix} & & \begin{pmatrix} 2 \\ 2 \end{pmatrix} \end{array}$$

$$\begin{array}{ccccc} G_{\alpha\beta} & \cdot & T_{k\delta} & = & W_{\alpha\beta k\delta} \\ \begin{pmatrix} 0 \\ 2 \end{pmatrix} & & \begin{pmatrix} 0 \\ 2 \end{pmatrix} & & \begin{pmatrix} 0 \\ 4 \end{pmatrix} \end{array}$$

3) We can contract a pair of indices one of which is up, the other down. The result is a tensor of type $\begin{pmatrix} N-1 \\ N'-1 \end{pmatrix}$

$$G^{\alpha\beta\gamma}{}_{\alpha ki} = G^{0\beta\gamma}{}_{0ki} + G^{1\beta\gamma}{}_{1ki} + G^{2\beta\gamma}{}_{2ki} + ... + G^{n\beta\gamma}{}_{nki} = G^{\beta\gamma}{}_{ki}$$

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

## 4.3 The metric Tensor

In chapter I we have seen that the metric tensor occupies 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* $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ *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)* $\vec{A} \cdot \vec{B}$

$$g(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B}. \tag{4.30}$$

The scalar product is usually defined to be a linear function of two vectors that satisfies the following properties

$$\begin{aligned} &\vec{U} \cdot \vec{V} = \vec{V} \cdot \vec{U} \\ &(a\vec{U}) \cdot \vec{V} = a(\vec{U} \cdot \vec{V}) \\ &(\vec{U} + \vec{V}) \cdot \vec{W} = \vec{U} \cdot \vec{W} + \vec{V} \cdot \vec{W} \end{aligned} \tag{4.31}$$

From eq. (4.31) it follows that **g** is a symmetric tensor. In fact

$$\vec{U} \cdot \vec{V} = g(\vec{U}, \vec{V}) = \vec{V} \cdot \vec{U} = g(\vec{V}, \vec{U}), \quad \rightarrow \quad g(\vec{U}, \vec{V}) = g(\vec{V}, \vec{U}). \tag{4.32}$$

Eqs. (4.31) and (4.31) 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

$$g_{\alpha\beta} = g(\vec{e}_{(\alpha)}, \vec{e}_{(\beta)}). \tag{4.33}$$

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

$$\vec{A} \cdot \vec{B} = g(\vec{A}, \vec{B}) \quad = \quad g(A^\alpha \vec{e}_{(\alpha)}, B^\beta \vec{e}_{(\beta)}) = A^\alpha B^\beta g(\vec{e}_{(\alpha)}, \vec{e}_{(\beta)}) = \tag{4.34}$$

$$A^\alpha B^\beta g_{\alpha\beta}.$$

---

EXAMPLES

1)

We want to show that the metric of Minkowski spacetime is

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

In fact, given the basis vector

$$\begin{aligned} \vec{e}_{(0)} &= \vec{e}_t \to (1,0,0,0) \\ \vec{e}_{(1)} &= \vec{e}_x \to (0,1,0,0) \\ \vec{e}_{(2)} &= \vec{e}_y \to (0,0,1,0) \\ \vec{e}_{(3)} &= \vec{e}_z \to (0,0,0,1) \end{aligned}$$

we know that they are mutually orthogonal, i.e.

$$\vec{e}_{(\alpha)} \cdot \vec{e}_{(\beta)} = 0 \qquad \text{if} \qquad \alpha \neq \beta.$$

It follows that

$$g_{(\alpha\beta)} = 0 \qquad \text{if} \qquad \alpha \neq \beta.$$

In addition the basis vector are unit vectors, i.e.

$$\vec{e}_{(k)} \cdot \vec{e}_{(k)} = 1 \qquad \text{if} \qquad k = 1, 3,$$

and

$$\vec{e}_{(0)} \cdot \vec{e}_{(0)} = -1$$

since $\vec{e}_{(0)}$ is a timelike vector. Consequently

$$g_{ik} = 1, \qquad \text{and} \qquad g_{00} = -1.$$

q.e.d.

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

2)

We now want to compute how the components of the metric tensor change if we change the coordinate system. We shall answer this question in two ways

a) by using the expressions of the new basis vectors and eq. (4.33)

b) by using the transformation law (4.26).

a)

For simplicity let us consider, as we did in the example in section 3.3, a 3-dimensional spacetime, and suppress the coordinate z.

$$\begin{aligned} \vec{e}_{(0)} &\rightarrow (1, 0, 0) \\ \vec{e}_{(1)} &\rightarrow (0, 1, 0) \\ \vec{e}_{(2)} &\rightarrow (0, 0, 1) \end{aligned}$$

$$
g_{ik} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & +1 \end{pmatrix} \equiv \eta_{ik} \quad i = k = 0, 2,
$$

and

$$
\begin{aligned}
\vec{e}_{(i)} \cdot \vec{e}_{(k)} &= 0 \quad if \;\; i \neq k \\
\vec{e}_{(0)} \cdot \vec{e}_{(0)} &= -1 \\
\vec{e}_{(1)} \cdot \vec{e}_{(1)} &= +1 = \vec{e}_{(2)} \cdot \vec{e}_{(2)}.
\end{aligned}
$$

We now change to polar coordinate

$$
x^0 = x^{0'}, \qquad x^1 = r\cos\theta, \qquad x^2 = r\sin\theta. \tag{4.35}
$$

The new basis vectors computed in sec. 3.3 are

$$
\vec{e}_{(0')} = \vec{e}_{(0)} \tag{4.36}
$$

$$
\vec{e}_{(1')} = \vec{e}_{(r)} = \cos\theta\vec{e}_{(1)} + \sin\theta\vec{e}_{(2)} \tag{4.37}
$$

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

Consequently

$$
\begin{aligned}
g_{0'0'} &= \vec{e}_{(0')} \cdot \vec{e}_{(0')} = \vec{e}_{(0)} \cdot \vec{e}_{(0)} = -1 \\
g_{0'i'} &= 0 \quad i' = 1, 2 \\
g_{1'1'} &= \vec{e}_{(1')} \cdot \vec{e}_{(1')} = (\cos\theta\vec{e}_{(1)} + \sin\theta\vec{e}_{(2)}) \cdot (\cos\theta\vec{e}_{(1)} + \sin\theta\vec{e}_{(2)}) = \cos^2\theta + \sin^2\theta = 1 \\
g_{2'2'} &= \vec{e}_{(2')} \cdot \vec{e}_{(2')} = +r^2\sin^2\theta + r^2\cos^2\theta = r^2 \\
g_{1'2'} &= -r\cos\theta\sin\theta + r\cos\theta\sin\theta = 0
\end{aligned}
$$

$$g_{i'k'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & r^2 \end{pmatrix}. \tag{4.38}$$

b) Let us now use the transformation law

$$g_{i'k'} = \Lambda^{\alpha}{}_{i'}\Lambda^{\beta}{}_{k'}g_{ik} = \Lambda^{\alpha}{}_{i'}\Lambda^{\beta}{}_{k'}\eta_{\alpha\beta}$$

Since $\eta_{ik}$ is diagonal we only need to consider $\alpha = \beta$.

$$g_{0'0'} = \Lambda^{\alpha}{}_{0'}\Lambda^{\alpha}{}_{0'}\eta_{\alpha\alpha} = \left(\frac{\partial x^{\alpha}}{\partial x^{0'}}\right)^2 \eta_{\alpha\alpha} = \left(\frac{\partial x^{0}}{\partial x^{0'}}\right)^2 \eta_{00} = 1 \cdot (-1) = -1$$

$$g_{0'i'} = \Lambda^{\alpha}{}_{0'}\Lambda^{\alpha}{}_{i'}\eta_{\alpha\alpha} = \frac{\partial x^0}{\partial x^{0'}}\frac{\partial x^0}{\partial x^{i'}}\eta_{00} + \frac{\partial x^1}{\partial x^{0'}}\frac{\partial x^1}{\partial x^{i'}}\eta_{11} + \frac{\partial x^2}{\partial x^{0'}}\frac{\partial x^2}{\partial x^{i'}}\eta_{22} = 0 \quad i' = 1, 2$$

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

$$\begin{aligned} g_{1'1'} &= \Lambda^{\alpha}{}_{1'}\Lambda^{\alpha}{}_{1'}\eta_{\alpha\alpha} = (\Lambda^{0}{}_{1'})^2\eta_{00} + (\Lambda^{1}{}_{1'})^2\eta_{11} + (\Lambda^{2}{}_{1'})^2\eta_{22} = \\ &= \left(\frac{\partial x^0}{\partial x^{1'}}\right)^2 \cdot (-1) + \left(\frac{\partial x^1}{\partial x^{1'}}\right)^2 \cdot 1 + \left(\frac{\partial x^2}{\partial x^{1'}}\right)^2 \cdot 1 = \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 \end{aligned}$$

$$g_{1'1'} = \cos^2\theta + \sin^2\theta = 1$$

Proceeding in this way we clearly find again the metric (4.38) in the new frame $(x^{0'}, x^{1'}, x^{2'}) = (ct, r, \theta)$.

### 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 a point $P(x^0, x^1, x^2)$ and $P'(x^0 + dx^0, x^1 + dx^1, x^2 + dx^2)$ is

$$\vec{ds} = dx^0 \vec{e}_{(0)} + dx^1 \vec{e}_{(1)} + dx^2 \vec{e}_{(2)} = dx^\alpha \vec{e}_{(\alpha)} \tag{4.39}$$

and $\vec{e}_{(\alpha)}$ are the basis vectors. By definition the metric tensor acting on two vectors produces their scalar product therefore

$$g(\vec{ds}, \vec{ds}) = \vec{ds} \cdot \vec{ds} = ds^2, \tag{4.40}$$

where $ds^2$ is the norm of the vector $\vec{ds}$, i.e. the square of the distance between $P$ and $P'$.

Eq. (4.40) gives

$$\begin{aligned} g(\vec{ds}, \vec{ds}) &= g(dx^0 \vec{e}_{(0)} + dx^1 \vec{e}_{(1)} + dx^2 \vec{e}_{(2)}, dx^0 \vec{e}_{(0)} + dx^1 \vec{e}_{(1)} + dx^2 \vec{e}_{(2)}) = \\ &= (dx^0)^2 g(\vec{e}_{(0)}, \vec{e}_{(0)}) + dx^1 dx^0 g(\vec{e}_{(1)}, \vec{e}_{(0)}) + dx^2 dx^0 g(\vec{e}_{(2)}, \vec{e}_{(0)}) + \\ &+ dx^0 dx^1 g(\vec{e}_{(0)}, \vec{e}_{(1)}) + (dx^1)^2 g(\vec{e}_{(1)}, \vec{e}_{(1)}) + dx^2 dx^1 g(\vec{e}_{(2)}, \vec{e}_{(1)}) + \\ &+ dx^0 dx^2 g(\vec{e}_{(0)}, \vec{e}_{(2)}) + dx^2 dx^1 g(\vec{e}_{(2)}, \vec{e}_{(1)}) + (dx^2)^2 g(\vec{e}_{(2)}, \vec{e}_{(2)}) \end{aligned}$$

We now remember that, according to eq. (4.33) if we apply $g$ to the basis vectors we get its components $g_{\alpha\beta}$, therefore the previous expression becomes

$$g(\vec{ds}, \vec{ds}) = (dx^0)^2 g_{00} + 2dx^0 dx^1 g_{01} + 2dx^0 dx^2 g_{02} + 2dx^1 dx^2 g_{12} + (dx^1)^2 g_{11} + (dx^2)^2 g_{22} \tag{4.41}$$

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

This calculation is simplified if we use the following notation

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

$$= \quad dx^{\alpha}dx^{\beta}g(\vec{e}_{(\alpha)},\vec{e}_{(\beta)}) = g_{\alpha\beta}dx^{\alpha}dx^{\beta} \tag{4.42}$$

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

This way of writing is completely equivalent to eq. (4.41). Thus, coming back to eq. (4.40) we find

$$ds^2 = g_{\alpha\beta}dx^{\alpha}dx^{\beta}. \tag{4.43}$$

For example, if the space is Minkowski spacetime $g_{\alpha\beta} = \eta_{\alpha\beta} = diag(-1,1,1)$, and eq. (4.43) gives

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2, \tag{4.44}$$

as expected.

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

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

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

$$ds^2 = g_{\alpha'\beta'}dx^{\alpha'}dx^{\beta'} \tag{4.45}$$

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

$$ds^2 = (dx^{0'})^2 g_{0'0'} + (dx^{1'})^2 g_{1'1'} + (dx^{2'})^2 g_{2'2'} = -(dx^0)^2 + dr^2 + r^2 d\theta^2. \tag{4.46}$$

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

### 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 $g(\ ,\vec{V})$, namely we leave the first slot empty. What is this? We know that if we fill the first slot with a generic vector $\vec{A}$ we will get a number, thus $g(\ ,\vec{V})$ 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 $g(\ ,\vec{V})$ is a one-form.

In addition, it is a particular one-form because it depends on $\vec{V}$: if I change $\vec{V}$, the one-form will be different. Let us indicate this one-form as

$$g(\ ,\vec{V}) = \tilde{V}(\ ). \tag{4.47}$$

We leave the slot empty because we can put any arbitrary vector into it.

By definition the components of $\tilde{V}$ are slot

$$V_\alpha = \tilde{V}(\vec{e}_{(\alpha)}) = g(\vec{e}_{(\alpha)},\vec{V}) = g(\vec{e}_{(\alpha)},V^\beta\vec{e}_{(\beta)}) = V^\beta g(\vec{e}_{(\alpha)},\vec{e}_{(\beta)}) = V^\beta g_{\alpha\beta},$$

hence

$$V_\alpha = g_{\alpha\beta} V^\beta. \tag{4.48}$$

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

In addition, if we multiply eq. (4.48) $g^{\alpha\gamma}$, where is the matrix inverse to $g_{\alpha\gamma}$

$$g_{\alpha\gamma} g^{\gamma\beta} = \delta_\alpha^\beta, \tag{4.49}$$

we find

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

i.e.

$$V^\gamma = g^{\alpha\gamma} V_\alpha, \tag{4.50}$$

Consequently the metric tensor also maps vectors into one-forms . 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

$$A_j^i = g_{jk} A^{ik},$$

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

$$A_{jl} = g_{jk} g_{ln} A^{kn},$$

or viceversa

$$A^{jl} = g^{jk} g^{ln} A_{kn}.$$

These maps are called index raising and lowering.

Summarizing, the metric tensor

1) allows to compute the inner product of two vectors $g(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B}$, 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 $ds^2 = g(\vec{ds}, \vec{ds}) = g_{\alpha\beta} dx^\alpha dx^\beta$.

3) It maps one-forms into vectors and viceversa.

4) It allows to raise and lower indices.

# Chapter 5

# Affine Connections

In chapter I 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 IV 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 I.

## 5.1 The covariant derivative of vectors

Be

$$\vec{V} = V^{\alpha}\vec{e}_{(\alpha)} \tag{5.1}$$

a vector. Is the derivative of $\vec{V}$, $\frac{\partial \vec{V}}{\partial x^{\beta}}$, a vector?

$$\frac{\partial \vec{V}}{\partial x^{\beta}} = \frac{\partial V^{\alpha}}{\partial x^{\beta}}\vec{e}_{(\alpha)} + V^{\alpha}\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}}. \tag{5.2}$$

The first term on the R.H.S. is a linear combination of the basis vectors $\{\vec{e}_{(\alpha)}\}$, therefore it is a vector. The term $\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}}$, is also a vector for the following reason. We can always make (locally) a coordinate transformation which brings the metric in the form $\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, and introduce new, constant basis vectors $\vec{e}_{(\alpha')}$ related to the old basis vectors $\vec{e}_{(\alpha)}$ by the equation

$$\vec{e}_{(\alpha)} = \Lambda^{\alpha'}{}_{\alpha}\vec{e}_{(\alpha')}. \tag{5.3}$$

Consequently

$$\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}} = \left(\frac{\partial}{\partial x^{\beta}}\Lambda^{\alpha'}{}_{\alpha}\right)\vec{e}_{(\alpha')},$$

and since the R.H.S. is a linear combination of vectors, $\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}}$ is a vector. Q.E.D.

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.:

$$\frac{\partial \vec{e}_{(\alpha)}}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta} \vec{e}_{(\mu)}, \tag{5.4}$$

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**.

Thus, the derivative of $\vec{V}$ in eq. (5.2) 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

$$\frac{\partial \vec{V}}{\partial x^\beta} = \left[ \frac{\partial V^\alpha}{\partial x^\beta} + V^\sigma \Gamma^\alpha_{\beta\sigma} \right] \vec{e}_{(\alpha)}. \tag{5.5}$$

For any fixed $\beta$, $\frac{\partial \vec{V}}{\partial x^\beta}$ is a vector field because, given a vector field $\vec{V}$, it produces at any point a new vector that 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

$$V^\alpha{}_{,\beta} = \frac{\partial V^\alpha}{\partial x} \quad , \qquad \text{and} \quad V^\alpha{}_{;\beta} = \frac{\partial V^\alpha}{\partial x^\beta} + V^\mu \Gamma^\alpha_{\beta\mu}, \tag{5.6}$$

eq. (5.5) becomes

$$\frac{\partial \vec{V}}{\partial x^\beta} = V^\alpha{}_{;\beta} \vec{e}_{(\alpha)}. \tag{5.7}$$

The quantities $V^{\alpha}{}_{;\beta}$ are the components of a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor. In fact, it is easy to show that a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor maps vectors into vectors and one-forms into one-forms.

Let us see why.

---

$T(\tilde{\omega}, \vec{v})$ is a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor field. Therefore it takes a one-form $\tilde{\omega}$, a vector $\vec{v}$ and associate a number to them.

If we write $T(\ \ , \vec{v})$, that means that if we fill the empty slot with any one-form we will get a number. Thus $T(\ \ , \vec{v})$ must be a vector (it follows from the definition of vectors). We shall call it

$$T(\ \ , \vec{v}) = \vec{T}(\ \ )$$

Let us find its components

$$\vec{T}(\tilde{\omega}^{(\alpha)}) = T^{\alpha} = T(\omega^{(\alpha)}, \vec{e}_{(\beta)} v^{\beta}) = v^{\beta} T(\tilde{\omega}^{(\alpha)}, \vec{e}_{(\beta)}) = v^{\beta} T^{\alpha}{}_{\beta}$$

$$T^{\alpha} = v^{\beta} T^{\alpha}{}_{\beta}$$

Similarly $T(\tilde{\sigma},\ \ ) = \tilde{T}(\ \ )$ is a one-form

$$\tilde{T}(\vec{e}_{(\alpha)}) = T_{\alpha} = T(\tilde{\sigma}, \vec{e}_{(\alpha)}) = T(\sigma_{\beta} \tilde{\omega}^{(\beta)}, \vec{e}_{(\beta)}) = \sigma_{\beta} T^{\beta}{}_{\alpha}$$

$$T_{\alpha} = \sigma_{\beta} T^{\beta}{}_{\alpha}$$

Thus we have shown that a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor maps vectors into vectors and one-forms into one-forms.

---

From eq.(5.7) it is clear that the $V^{\alpha}{}_{;\beta}$ can be seen as the components of a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ tensor which maps the vector $\vec{e}_{(\alpha)}$ into the vector $\frac{\partial \vec{V}}{\partial x^{\beta}}$. This tensor field is called **Covariant derivative of a vector**, it is denoted as $\nabla\vec{V}$, and

$$(\nabla\vec{V})^{\alpha}{}_{\beta} \equiv \nabla_{\beta}V^{\alpha} = V^{\alpha}{}_{;\beta}. \tag{5.8}$$

In a locally inertial frame the basis vectors are constant, and consequently, according to eq. (5.4) the affine connections vanish and from eq. (5.6) it follows that

$$V^{\alpha}{}_{;\beta} = V^{\alpha}{}_{,\beta} \Longrightarrow \frac{\partial \vec{V}}{\partial x^{\beta}} = V^{\alpha}{}_{,\beta}\vec{e}_{(\alpha)}. \tag{5.9}$$

**In a locally inertial frame covariant and ordinary derivative coincide.**

In this chapter we have now introduced the connections as those quantities that allow to find the covariant derivative of a vector in an arbitrary frame. How can we compute $\Gamma^{\mu}_{\alpha\beta}$ from what we know? Let us consider for example a 2-dimensional flat space in polar coordinates, and remember that the equation which defines $\Gamma^{\mu}_{\alpha\beta}$ is

$$\frac{\partial \vec{e}_{(\alpha)}}{\partial x^{\beta}} = \Gamma^{\mu}_{\alpha\beta}\vec{e}_{(\mu)}. \tag{5.10}$$

Then

$$\frac{\partial \vec{e}_{(1)}}{\partial x^1} = \frac{\partial \vec{e}_r}{\partial r} = \frac{\partial}{\partial r}(\cos\theta \vec{e}_x + \sin\theta \vec{e}_y) = 0,$$

and consequently

$$\Gamma^{\mu}_{rr}\vec{e}_{(\mu)} = \Gamma^{r}_{rr}\vec{e}_r + \Gamma^{\theta}_{rr}\vec{e}_{(\theta)} = 0 \Longrightarrow \Gamma^{r}_{rr} = \Gamma^{\theta}_{rr} = 0.$$

Moreover

$$\begin{aligned}\frac{\partial \vec{e}_{(1)}}{\partial x^2} &= \frac{\partial \vec{e}_r}{\partial \theta} = \frac{\partial}{\partial \theta}(\cos\theta \vec{e}_x + \sin\theta \vec{e}_y) = \\ &= -\sin\theta \vec{e}_x + \cos\theta \vec{e}_y = \frac{1}{r}\vec{e}_\theta,\end{aligned}$$

therefore

$$\frac{1}{r}\vec{e}_\theta = \Gamma^{\mu}_{r\theta}\vec{e}_{(\mu)} = \Gamma^{r}_{r\theta}\vec{e}_r + \Gamma^{\theta}_{r\theta}\vec{e}_{(\theta)} \Longrightarrow \Gamma^{r}_{r\theta} = 0 \ \ , \ \ \Gamma^{\theta}_{r\theta} = \frac{1}{r}.$$

Proceeding along these lines one can show that

$$\Gamma^{r}_{\theta r} = 0 \ , \ \Gamma^{\theta}_{\theta r} = \frac{1}{r} \ , \ \Gamma^{r}_{\theta\theta} = -r \ , \ \Gamma^{\theta}_{\theta\theta} = 0.$$

It should be noted that altough 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 $\theta$.

## 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

$$\nabla_\mu \Phi = \frac{\partial \Phi}{\partial x^\mu} = (\tilde{d}\Phi)_\mu. \tag{5.11}$$

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

$$\tilde{q}(\vec{V}) = q_j V^j, \tag{5.12}$$

where $q_j$ and $V^j$ are the components of the one-form and of the vector. We can therefore assume that the scalar function $\Phi$ is

$$\Phi = q_j V^j, \tag{5.13}$$

and consequently the covariant derivative will be

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

Substituting $\frac{\partial V^j}{\partial x^\mu}$ from eq. (5.6) we find

$$\nabla_\mu \Phi = \frac{\partial q_j}{\partial x^\mu} V^j + q_j [V^j{}_{;\mu} - V^k \Gamma^j_{\mu k}].$$

We relabel the indices to put $V^j$ in evidence

$$\begin{aligned}\nabla_\mu \Phi &= \frac{\partial q_j}{\partial x^\mu} V^j + q_k V^k{}_{;\mu} - q_k V^j \Gamma^k_{\mu j} = \\ &= [\frac{\partial q_j}{\partial x^\mu} - q_k \Gamma^k_{\mu j}] V^j + q_k V^k{}_{;\mu}. \end{aligned} \tag{5.14}$$

$\nabla_\mu \Phi$ are the components of a tensor ( $\nabla_\mu \Phi \equiv \tilde{df}_\mu$) as well as $V^j$, $q_k$ and $V^k{}_{;\mu}$ [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

$$(\nabla \tilde{q})_{\alpha\beta} \equiv \nabla_\beta \tilde{q}_\alpha = q_{\alpha;\beta} = q_{\alpha,\beta} - q_k \Gamma^k_{\beta\alpha}. \tag{5.15}$$

We now observe that eq. (5.14) can be written as

$$\nabla_\mu \Phi = \nabla_\mu (q_\alpha V^\alpha) = q_{\alpha;\mu} V^\alpha + q_\alpha V^\alpha{}_{;\mu}. \tag{5.16}$$

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)

$$(\nabla T_{\mu\nu})_\beta = T_{\mu\nu,\beta} - T_{\alpha\nu} \Gamma^\alpha_{\beta\mu} - T_{\mu\alpha} \Gamma^\alpha_{\beta\nu} \tag{5.17}$$

[1] here we use the word 'tensor' to indicate any type of tensors, included vectors and one-forms

$$(\nabla A^{\mu\nu})_\beta = A^{\mu\nu}{}_{,\beta} + A^{\alpha\nu}\Gamma^\mu_{\alpha\beta} + A^{\mu\alpha}\Gamma^\nu_{\alpha\beta} \tag{5.18}$$

$$(\nabla B^\mu{}_\nu)_\beta = B^\mu{}_{\nu,\beta} + B^\alpha{}_\nu\Gamma^\mu_{\beta\alpha} - B^\mu_\alpha\Gamma^\alpha_{\beta\nu} \tag{5.19}$$

what is the rule?

## 5.3 The covariant derivative of the metric tensor

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

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

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. 5.4). In this frame

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

$g_{\alpha\beta;\mu}$ 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

$$g_{\alpha\beta;\mu} = 0 \tag{5.20}$$

**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:

$$\Phi_{,\beta;\alpha} = \Phi_{,\beta,\alpha} = \frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\Phi, \tag{5.21}$$

and since partial derivatives commute

$$\Phi_{,\beta,\alpha} = \Phi_{\alpha,\beta} \Rightarrow \Phi_{,\beta;\alpha} = \Phi_{,\alpha;\beta}. \tag{5.22}$$

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

$$\Phi_{,\beta,\alpha} - \Phi_{,\mu}\Gamma^\mu_{\beta\alpha} = \Phi_{\alpha,\beta} - \Phi_{,\mu}\Gamma^\mu_{\alpha\beta}$$

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

$$\Phi_{,\mu}\Gamma^\mu_{\beta\alpha} = \Phi_{,\mu}\Gamma^\mu_{\alpha\beta},$$

and consequently

$$\Gamma^\mu_{\beta\alpha} = \Gamma^\mu_{\alpha\beta} \tag{5.23}$$

in any coordinate system. Q.E.D.

## 5.5 The relation between the affine connections and the metric tensor

From eq. (5.20) 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

$$g_{\alpha\beta,\mu} = \Gamma^{\nu}_{\alpha\mu} g_{\nu\beta} + \Gamma^{\nu}_{\beta\mu} g_{\alpha\nu}. \tag{5.24}$$

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

$$\begin{aligned} g_{\alpha\beta,\mu} \quad &+ \quad g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = (\Gamma^{\nu}_{\alpha\mu} - \Gamma^{\nu}_{\mu\alpha}) g_{\nu\beta} + \\ &+ \quad (\Gamma^{\nu}_{\beta\mu} + \Gamma^{\nu}_{\mu\beta}) g_{\alpha\nu} + (\Gamma^{\nu}_{\alpha\beta} - \Gamma^{\nu}_{\beta\alpha}) g_{\nu\mu}, \end{aligned}$$

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

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

$$g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = 2\Gamma^{\nu}_{\beta\mu} g_{\alpha\nu}.$$

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 the expression of the affine connections in terms of the metric

$$\Gamma^{\gamma}_{\beta\mu} = \frac{1}{2} g^{\alpha\gamma}(g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}) \tag{5.25}$$

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

$$\Gamma^{\alpha}_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial \xi^{\lambda}} \frac{\partial^2 \xi^{\lambda}}{\partial x^{\mu} \partial x^{\nu}}. \tag{5.26}$$

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

$$\frac{d^2 \xi^{\alpha}}{d\tau^2} = 0. \tag{5.27}$$

Then we transformed this equation to an arbitrary coordinate system $\{x^{\alpha}\}$ and we showed that it becomes

$$\frac{d^2 x^{\alpha}}{d\tau^2} + \Gamma^{\alpha}_{\mu\nu} \left[ \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau} \right] = 0, \tag{5.28}$$

with $\Gamma^{\alpha}_{\mu\nu}$ defined in eq. (5.26).

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

$$\frac{\partial \vec{e}_{(\mu)}}{\partial x^{\nu}} = \Gamma^{\alpha}_{\mu\nu} \vec{e}_{(\alpha)}. \tag{5.29}$$

What is the relation between eq. (5.26) and eq. (5.29)?

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

$$\vec{e}_{(\mu')} = \Lambda^\alpha{}_{\mu'}\vec{e}_{M(\alpha)} = \frac{\partial \xi^\alpha}{\partial x^{\mu'}}\vec{e}_{M(\alpha)}. \tag{5.30}$$

In this frame, eq. (5.29)which defines the affine connections can be rewritten as

$$\frac{\partial}{\partial x^{\nu'}}\left[\Lambda^\beta{}_{\mu'}\vec{e}_{M(\beta)}\right] = \Gamma^{\alpha'}_{\mu'\nu'}\Lambda^\gamma{}_{\alpha'}\vec{e}_{M(\gamma)} \tag{5.31}$$

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

$$\frac{\partial \Lambda^\beta{}_{\mu'}}{\partial x^{\nu'}}\vec{e}_{M(\beta)} = \Gamma^{\alpha'}_{\mu'\nu'}\Lambda^\gamma{}_{\alpha'}\vec{e}_{M(\gamma)}. \tag{5.32}$$

This equation can be re-written as

$$\left(\frac{\partial \Lambda^\beta{}_{\mu'}}{\partial x^{\nu'}} - \Gamma^{\alpha'}_{\mu'\nu'}\Lambda^\beta{}_{\alpha'}\right)\vec{e}_{M(\beta)} = 0. \tag{5.33}$$

We now multiply eq. (5.33) by $\Lambda^{\sigma'}{}_\beta$ and find

$$\Lambda^{\sigma'}{}_\beta\frac{\partial \Lambda^\beta{}_{\mu'}}{\partial x^{\nu'}} - \Gamma^{\alpha'}_{\mu'\nu'}\Lambda^{\sigma'}{}_\beta\Lambda^\beta{}_{\alpha'} = 0. \tag{5.34}$$

Since $\Lambda^{\sigma'}{}_\beta\Lambda^\beta{}_{\alpha'} = \delta^{\sigma'}{}_{\alpha'}$, it follows that

$$\Gamma^{\sigma'}_{\mu'\nu'} = \Lambda^{\sigma'}{}_\beta\frac{\partial \Lambda^\beta{}_{\mu'}}{\partial x^{\nu'}} = \frac{\partial x^{\sigma'}}{\partial x^\beta}\frac{\partial^2 x^\beta}{\partial x^{\nu'}\partial x^{\mu'}},$$

which coincides with eq. (5.26). Thus, as axpected, the two definitions are equivalent. How do the $\Gamma^\alpha_{\beta\gamma}$ transform?

The easiest way to see it is from the definition (5.26). In an arbitrary coordinate system $\{x^{\mu'}\}$ they are

$$
\begin{aligned}
\Gamma^{\lambda'}_{\mu'\nu'} &= \frac{\partial x^{\lambda'}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\nu'} \partial x^{\mu'}} = \\
&= \frac{\partial x^{\lambda'}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial}{\partial x^{\mu'}} \left( \frac{\partial \xi^{\alpha}}{\partial x^{\sigma}} \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \right) = \\
&= \frac{\partial x^{\lambda'}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \left[ \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\tau} \partial x^{\sigma}} \frac{\partial x^{\tau}}{\partial x^{\mu'}} + \frac{\partial \xi^{\alpha}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\mu'}} \right] = \\
&= \frac{\partial x^{\lambda'}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x^{\nu'}} \frac{\partial x^{\tau}}{\partial x^{\mu'}} \Gamma^{\rho}_{\tau\sigma} + \frac{\partial x^{\lambda'}}{\partial x^{\sigma}} \frac{\partial^2 x^{\sigma}}{\partial x^{\nu'} \partial x^{\mu'}}
\end{aligned} \tag{5.35}
$$

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

## 5.6 Appendix C non coordinate basis

In Chapter 3 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

$$
\{\vec{e}_{(\alpha)}\} \to \begin{cases} \vec{e}_{(0)} & \to \quad (1,0,0) \\ \vec{e}_{(1)} & \to \quad (0,1,0) \\ \vec{e}_{(2)} & \to \quad (0,0,1) \end{cases} \tag{5.36}
$$

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

$$
\begin{cases} \vec{e}_{(0')} = \vec{e}_{(0)} \\ \vec{e}_{(1')} = \vec{e}_r = \cos\theta \vec{e}_{(1)} + \sin\theta \vec{e}_{(2)} \\ \vec{e}_{(2')} = \vec{e}_r = -r\sin\theta \vec{e}_{(1)} + r\cos\theta \vec{e}_{(2)} \end{cases} \tag{5.37}
$$

according to the law

$$\vec{e}_{(\alpha')} = \Lambda^k{}_{\alpha'} \vec{e}_{(k)}.$$

The new basis is a coordinate basis and the matrix $\Lambda^k{}_{\alpha'} = \frac{\partial x^k}{\partial x^{\alpha'}}$ 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

$$\begin{cases} \vec{e}_{\hat{r}} = \vec{e}_r \\ \vec{e}_{\hat{t}} = \vec{e}_t \\ \vec{e}_{\hat{\theta}} = \frac{1}{r}\vec{e}_\theta. \end{cases} \tag{5.38}$$

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

$$e_{(\hat{\alpha})} = \Lambda^i{}_{\hat{\alpha}} \vec{e}_{(i)} = \frac{\partial x^i}{\partial x^{\hat{\alpha}}} \vec{e}_{(i)}$$

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 $\{\omega^{(\alpha')}\}$ 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.

$$\omega^{(\hat{\alpha})} = \Lambda^{\hat{\alpha}}{}_{\beta}\tilde{\omega}^{(\beta)} = \frac{\partial x^{\hat{\alpha}}}{\partial x^{\beta}}\tilde{\omega}^{(\beta)}?$$

(Show that in the previous example

$$\begin{aligned}\tilde{\omega}^{\hat{1}} &= \tilde{\omega}^{r} = \cos\theta\tilde{dx} + \sin\theta\tilde{dy} \\ \tilde{\omega}^{\hat{2}} &= \tilde{\omega}^{\theta} = r\tilde{d\theta} = -\sin\theta\tilde{dx} + \cos\theta\tilde{dy}\ )\end{aligned} \tag{5.39}$$

The point is that if this is true, $\Lambda^{\hat{\alpha}}{}_{\beta}$ **must coincide** with the partial derivative $\frac{\partial x^{\hat{\alpha}}}{\partial x^{\beta}}$, and consequently the following condition must be satisfied for any $\Lambda^{\hat{\alpha}}{}_{\gamma}$:

$$\frac{\partial}{\partial x^{\gamma}}\Lambda^{\hat{\alpha}}{}_{\beta} = \frac{\partial^2 x^{\hat{\alpha}}}{\partial x^{\gamma}\partial x^{\beta}} = \frac{\partial^2 x^{\hat{\alpha}}}{\partial x^{\beta}\partial x^{\gamma}} = \frac{\partial}{\partial x^{\beta}}\Lambda^{\hat{\alpha}}{}_{\gamma}. \tag{5.40}$$

This is an "integrability condition" that all the components of $\Lambda^{\hat{\alpha}}{}_{\gamma}$ must satisfy in order the coordinates $\{x^{\hat{\alpha}}\}$ do exist.

For examples, let us check whether the basis (5.39) is a coordinate basis. From the expression of $\tilde{\omega}^{\theta}$ we find that

$$\Lambda^{\hat{2}}{}_{1} = \frac{\partial x^{\hat{2}}}{\partial x} = -\sin\theta \qquad \Lambda^{\hat{2}}{}_{2} = \frac{\partial x^{\hat{2}}}{\partial y} = \cos\theta,$$

eq. (5.40) gives

$$\frac{\partial}{\partial y}\Lambda^{\hat{2}}{}_{1} = \frac{\partial}{\partial x}\Lambda^{\hat{2}}{}_{2} \ \Rightarrow\ \frac{\partial}{\partial y}(-\sin\theta) = \frac{\partial}{\partial x}(\cos\theta),$$

But

$$x = r\cos\theta \qquad y = r\sin\theta \qquad r = \sqrt{x^2+y^2},$$

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 $\{\tilde{\omega}^{(\hat{\alpha})}\}$ **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:

$$\tilde{d}\Phi \;\rightarrow\; \left\{\frac{\partial\Phi}{\partial x^\alpha}\right\} \equiv \{\Phi_{,\alpha}\} \;. \tag{5.41}$$

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:

$$\tilde{d}\Phi_{j'} = \Lambda^k{}_{j'}\tilde{d}\Phi_k$$

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

$$\begin{cases} \tilde{d}\Phi_r = \Lambda^x{}_r\tilde{d}\Phi_x + \Lambda^y{}_r\tilde{d}\Phi_y = \dfrac{\partial x}{\partial r}\tilde{d}\Phi_x + \dfrac{\partial y}{\partial r}\tilde{d}\Phi_y \\ \tilde{d}\Phi_\theta = \Lambda^x{}_\theta\tilde{d}\Phi_x + \Lambda^y{}_\theta\tilde{d}\Phi_y = \dfrac{\partial x}{\partial \theta}\tilde{d}\Phi_x + \dfrac{\partial y}{\partial \theta}\tilde{d}\Phi_y. \end{cases} \tag{5.42}$$

Being

$$x = r\cos\theta,$$

$$y = r\sin\theta\ ,$$

$$\begin{cases} \tilde{d}\Phi_r = \cos\theta\tilde{d}\Phi_x + \sin\theta\tilde{d}\Phi_y = \dfrac{\partial\Phi}{\partial r} = \Phi_{,r} \\ \tilde{d}\Phi_\theta = -r\sin\theta\tilde{d}\Phi_x + +r\cos\theta\tilde{d}\Phi_y = \dfrac{\partial\Phi}{\partial\theta} = \Phi_{,\theta}. \end{cases} \tag{5.43}$$

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

$$\tilde{d}\Phi_{\jmath'} \ \rightarrow \frac{\partial\Phi}{\partial x^{\jmath'}}\ .$$

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 $\tilde{d}\Phi_{\hat{\jmath}} = \frac{\partial\Phi}{\partial x^{\hat{\jmath}}}$ !

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

$$\nabla_\alpha\vec{e}_{(\beta)} = \frac{\partial\vec{e}_{(\beta)}}{\partial x^\alpha} = \Gamma^{\nu}_{\beta\alpha}\vec{e}_{(\nu)}\ . \tag{5.44}$$

This is a definition valid in any basis, therefore in terms of a noncoordinate basis $\{\vec{e}_{(\hat{\alpha})}\}$ eq. (5.44) becomes

$$\nabla_{\hat{\alpha}}\vec{e}_{(\hat{\beta})} = \Gamma^{\hat{\nu}}_{\hat{\beta}\hat{\alpha}}\vec{e}_{(\hat{\nu})}\ . \tag{5.45}$$

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.(5.22) we see that we used this condition to show the simmetry of the affine conection 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 (5.25) 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 explicitely specify when we will use a non coordinate basis.

# Chapter 6

# Parallel Transport

## 6.1 Summary of the preceeding chapters

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

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu}\left[\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\right] = 0. \tag{6.1}$$

In this frame the line element is

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu. \tag{6.2}$$

Then we have seen that the Equivalence Principle allows to find a locally inertial frame $\{\xi^\alpha\}$ where eq. (6.1) becomes

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0, \tag{6.3}$$

and the line element reduces to

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{6.4}$$

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, or the knowledge of the $\Gamma^\alpha_{\mu\nu}$ do not allow to decide whether we are in the presence of a 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

$$\vec{A} = A^\alpha \vec{e}_{(\alpha)}. \tag{6.5}$$

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

$$\nabla_\beta V^\alpha = V^\alpha{}_{,\beta} + \Gamma^\alpha{}_{\mu\beta} V^\mu. \tag{6.6}$$

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

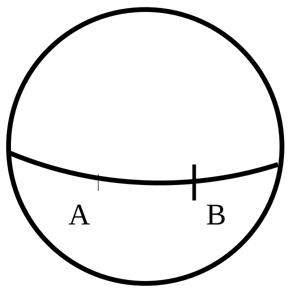
1) the spacetime is flat **and** we are in a basis where the vectors $\vec{e}_{(\alpha)}$ are constant. Consequently $\Gamma^{\alpha}{}_{\mu\beta} = 0$.

2) the spacetime is curved, but we are in a locally inertial frame. In fact, in this frame eq. (6.1) reduces to eq. (6.3), which means again that $\Gamma^{\alpha}{}_{\mu\beta} = 0$.

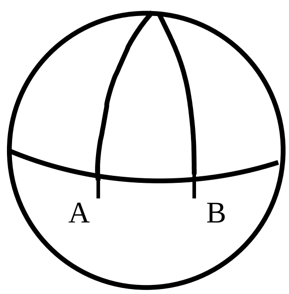
The fact that we can always find a frame where $g_{\mu\nu}$ reduces to $\eta_{\mu\nu}$ and the $\Gamma^{\alpha}{}_{\mu\beta} = 0$ (and consequently the first derivatives of $g_{\mu\nu}$ vanish) implies that in order to know if we are in the presence of a gravitational field, (i.e. if the spacetime is curved), **we need to know the second derivatives of the metric tensor** $g_{\mu\nu,\alpha,\beta}$. This result should not be surprising. In fact in chapter I when we introduced the 2-dimensional gaussian geometry 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 if the space is flat or curved. The gaussian curvature is computed from product of first derivatives and from the second derivatives of the metric, thus we are now looking 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? At the moment we only know that it should contain the second derivatives of $g_{\mu\nu}$. But this is not enough. In order to introduce the curvature tensor we first need to introduce the notion of parallel transport of a vector along a curve.

**Parallel Transport**

In the chapter I 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:



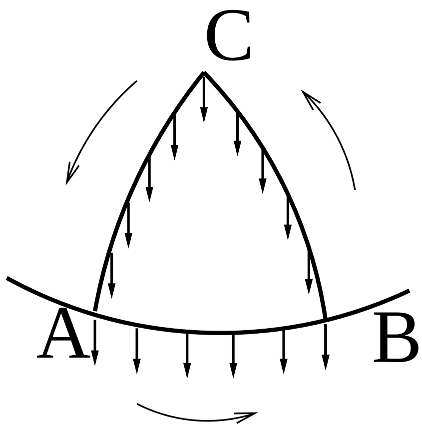

consider two segments in A and B, perpendicular to the equator, i.e. parallel.




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 lenght**. Let us consider first the case when the path belongs to a flat space.

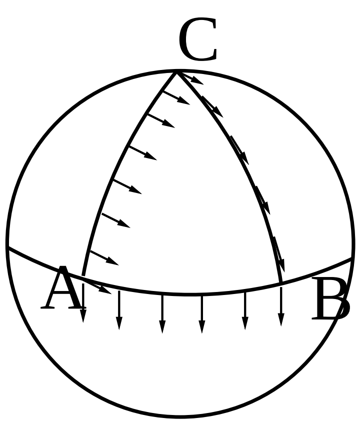
a) *FLAT SPACE*




When we return to A the displaced vector coincides with the original vector in A.

b) *ON A SPHERE*

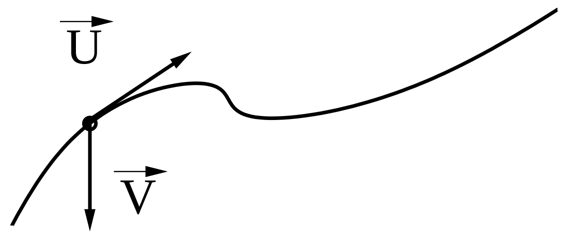
(remember that the vector must always be tangent to the sphere)




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 parallely transported. Consider a curve of parameter $\lambda$ and a vector field $\vec{V}$ defined at every point of the curve. Be $\vec{U} \rightarrow \{\frac{dx^\alpha}{d\lambda}\}$ the vector tangent to the curve




At every point of the curve we can choose a locally inertial frame. In this frame, if we move $\vec{V}$ along the curve of an infinitesimal $d\lambda$, parallel to itself and keeping its lenght unchanged, its components do not change

$$\frac{dV^\alpha}{d\lambda} = 0. \tag{6.7}$$

But

$$\frac{dV^\alpha}{d\lambda} = \frac{dV^\alpha}{dx^\beta}\frac{dx^\beta}{d\lambda} = U^\beta V^\alpha{}_{,\beta} = 0. \tag{6.8}$$

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

$$\frac{dV^\alpha}{d\lambda} = U^\beta V^\alpha{}_{;\beta} = 0. \tag{6.9}$$

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. (6.9) is the frame-invariant definition of the **parallel transport** of $\vec{V}$ along the curve identified by the tangent vector $\vec{U}$. We can therefore use the following equivalent notations

$$U^\beta V^\alpha{}_{;\beta} = 0, \qquad \to \frac{d\vec{V}}{d\lambda} = \nabla_{\vec{U}}\vec{V} = 0, \tag{6.10}$$

where

$$\left(\nabla_{\vec{U}}\vec{V}\right)^\alpha = \frac{dx^\beta}{d\lambda}\left[\frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha{}_{\beta\nu}V^\nu\right] = \frac{dV^\alpha}{d\lambda} + \Gamma^\alpha{}_{\beta\nu}V^\nu U^\beta. \tag{6.11}$$

From eq. (6.11) it follows that, contrary to what happens in flat space, the components of a vector parallel-transported along a curve in curved space do change.

## 6.2 Geodesics are those curves which parallel-transport their own tangent vectors

Let us prove this statement. It says that geodesics are those curves such that

$$\nabla_{\vec{U}}\vec{U} = 0. \tag{6.12}$$

In components this becomes

$$U^\beta U^\alpha{}_{;\beta} = U^\beta[U^\alpha{}_{,\beta} + \Gamma^\alpha{}_{\mu\beta}U^\mu] = 0. \tag{6.13}$$

If $\lambda$ is the parameter $U^\beta = \frac{dx^\beta}{d\lambda}$, and

$$U^\beta U^\alpha{}_{,\beta} = \left(\frac{\partial}{\partial x^\beta}\frac{dx^\alpha}{d\lambda}\right)\frac{dx^\beta}{d\lambda} = \frac{d^2x^\alpha}{d\lambda^2}; \tag{6.14}$$

eq. (6.13) becomes

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\mu\beta}\left[\frac{dx^\mu}{d\lambda}\frac{dx^\beta}{d\lambda}\right] = 0, \tag{6.15}$$

which is indeed the geodesic equation. The only difference is that in place of the proper time there is the parameter $\lambda$. However we can change the parameter (and make it coincides with $\tau$), which means that the path remains the same, but the curve changes. But there are some restrictions if we require the new curve to be a geodesic:

**If $\lambda$ is a parameter of a geodesic, only linear transformations of $\lambda$ :**

$$s = a\lambda + b, \qquad\qquad a, b = \textbf{const}, \tag{6.16}$$

**give new parameters in which the geodesic equation is satisfied.**

In fact

$$\frac{d}{d\lambda} = \frac{d}{ds}\frac{ds}{d\lambda} = a\frac{d}{ds}, \tag{6.17}$$

and it is immediate to see that eq. (6.15) becomes

$$\frac{d^2x^\alpha}{ds^2} + \Gamma^\alpha_{\mu\nu}\left[\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}\right] = 0. \tag{6.18}$$

$\lambda$ and $s$ are called **affine parameters**.

# Chapter 7

# The Curvature Tensor

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

## 7.1 a) A Formal Approach

Let us start writing the transformation rule for affine connections

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{\partial x^{\lambda}}{\partial x^{\tau\prime}} \frac{\partial x^{\rho\prime}}{\partial x^{\mu}} \frac{\partial x^{\sigma\prime}}{\partial x^{\nu}} \Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime} + \frac{\partial x^{\lambda}}{\partial x^{\tau\prime}} \frac{\partial^2 x^{\tau\prime}}{\partial x^{\mu} \partial x^{\nu}}. \tag{7.1}$$

As we already noticed (Chapter V sec. 4) 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. (7.1) by $\frac{\partial x^{\tau\prime}}{\partial x^{\lambda}}$:

$$\frac{\partial^2 x^{\tau\prime}}{\partial x^{\mu} \partial x^{\nu}} = \frac{\partial x^{\tau\prime}}{\partial x^{\lambda}} \Gamma^{\lambda}{}_{\mu\nu} - \frac{\partial x^{\rho\prime}}{\partial x^{\mu}} \frac{\partial x^{\sigma\prime}}{\partial x^{\nu}} \Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}. \tag{7.2}$$

We now differentiate this equation with respect to $x^k$

$$\frac{\partial^3 x^{\tau\prime}}{\partial x^k \partial x^\mu \partial x^\nu} = \frac{\partial^2 x^{\tau\prime}}{\partial x^k \partial x^\lambda}\Gamma^\lambda{}_{\mu\nu} + \frac{\partial x^{\tau\prime}}{\partial x^\lambda}\left(\frac{\partial}{\partial x^k}\Gamma^\lambda{}_{\mu\nu}\right) \tag{7.3}$$
$$-\frac{\partial^2 x^{\rho\prime}}{\partial x^k \partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime} - \frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial^2 x^{\sigma\prime}}{\partial x^k \partial x^\nu}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime} - \frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\left(\frac{\partial}{\partial x^k}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right).$$

We now use eq. (7.2):

$$\frac{\partial^3 x^{\tau\prime}}{\partial x^k \partial x^\mu \partial x^\nu} = \tag{7.4}$$
$$+\Gamma^\lambda{}_{\mu\nu}\left[\frac{\partial x^{\tau\prime}}{\partial x^\alpha}\Gamma^\alpha{}_{k\lambda} - \frac{\partial x^{i\prime}}{\partial x^k}\frac{\partial x^{j\prime}}{\partial x^\lambda}\Gamma^{\tau\prime}{}_{i\prime j\prime}\right] + \frac{\partial x^{\tau\prime}}{\partial x^\lambda}\left[\frac{\partial}{\partial x^k}\Gamma^\lambda{}_{\mu\mu}\right]$$
$$-\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\left[\frac{\partial x^{\rho\prime}}{\partial x^\alpha}\Gamma^\alpha{}_{k\mu} - \frac{\partial x^{i\prime}}{\partial x^k}\frac{\partial x^{j\prime}}{\partial x^\mu}\Gamma^{\rho\prime}{}_{i\prime j\prime}\right]$$
$$-\frac{\partial x^{\rho\prime}}{\partial x^\mu}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\left[\frac{\partial x^{\sigma\prime}}{\partial x^\alpha}\Gamma^\alpha{}_{k\nu} - \frac{\partial x^{i\prime}}{\partial x^k}\frac{\partial x^{j\prime}}{\partial x^\nu}\Gamma^{\sigma\prime}{}_{i\prime j\prime}\right]$$
$$-\frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\left(\frac{\partial}{\partial x^k}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right).$$

Let us rewrite the last term as

$$\frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\frac{\partial x^{\eta\prime}}{\partial x^k}\left(\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right). \tag{7.5}$$

(The reason is that the indices of $\Gamma$ have a prime, thus the derivatives must be computed with respect to the $\{x^{\alpha\prime}\}$). We now rewrite eq. (7.5) in the following way

$$\frac{\partial^3 x^{\tau\prime}}{\partial x^k \partial x^\mu \partial x^\nu} = \tag{7.6}$$

$$\left[\frac{\partial x^{\tau\prime}}{\partial x^{\lambda}}\left(\frac{\partial}{\partial x^{k}}\Gamma^{\lambda}{}_{\mu\nu}\right)+\left(\frac{\partial x^{\tau\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\lambda}\Gamma^{\lambda}{}_{\mu\nu}\right)\right]$$
$$-\left[\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\frac{\partial x^{\eta\prime}}{\partial x^{k}}\left(\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right)-\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\Gamma^{\rho\prime}{}_{i\prime j\prime}\right]$$
$$-\left[\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\Gamma^{\sigma\prime}{}_{i\prime j\prime}\right]$$
$$-\left[\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\frac{\partial x^{\rho\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\mu}+\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\frac{\partial x^{\sigma\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\nu}+\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{\mu\mu}\Gamma^{\tau\prime}{}_{i\prime j\prime}\right].$$

We now relabel the indices in the following way

$$\frac{\partial x^{\tau\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\lambda}\Gamma^{\lambda}{}_{\mu\nu}\rightarrow\frac{\partial x^{\tau\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{k\eta}\Gamma^{\eta}{}_{\mu\nu}\tag{7.7}$$

$$\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\Gamma^{\rho\prime}{}_{i\prime j\prime}\rightarrow\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\frac{\partial x^{\eta\prime}}{\partial x^{k}}\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\lambda\prime\sigma\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\rho\prime}$$
$$\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\Gamma^{\sigma\prime}{}_{i\prime j\prime}\rightarrow\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\frac{\partial x^{\eta\prime}}{\partial x^{k}}\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\rho\prime\lambda\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\sigma\prime}$$

$$\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\frac{\partial x^{\rho\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\mu}\rightarrow\frac{\partial x^{\rho\prime}}{\partial x^{\nu}}\Gamma^{\tau\prime}{}_{\sigma\prime\rho\prime}\frac{\partial x^{\sigma\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{k\mu}$$
$$\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\frac{\partial x^{\sigma\prime}}{\partial x^{\alpha}}\Gamma^{\alpha}{}_{k\nu}\rightarrow\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\frac{\partial x^{\sigma\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{k\nu}$$
$$\frac{\partial x^{i\prime}}{\partial x^{k}}\frac{\partial x^{j\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{\mu\mu}\Gamma^{\tau\prime}{}_{i\prime j\prime}\rightarrow\frac{\partial x^{\rho\prime}}{\partial x^{k}}\frac{\partial x^{\sigma\prime}}{\partial x^{\lambda}}\Gamma^{\lambda}{}_{\mu\mu}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}$$

With these changes the terms can be collected in the following way

$$\frac{\partial^{3}x^{\tau\prime}}{\partial x^{k}\partial x^{\mu}\partial x^{\nu}}=\frac{\partial x^{\tau\prime}}{\partial x^{\lambda}}\left[\left(\frac{\partial}{\partial x^{k}}\Gamma^{\lambda}{}_{\mu\nu}\right)+\Gamma^{\lambda}{}_{k\eta}\Gamma^{\eta}{}_{\mu\nu}\right]\tag{7.8}$$
$$-\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}\frac{\partial x^{\sigma\prime}}{\partial x^{\nu}}\frac{\partial x^{\eta\prime}}{\partial x^{k}}\left[\left(\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right)-\Gamma^{\tau\prime}{}_{\lambda\prime\sigma\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\rho\prime}-\Gamma^{\tau\prime}{}_{\rho\prime\lambda\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\sigma\prime}\right]$$
$$-\frac{\partial x^{\sigma\prime}}{\partial x^{\lambda}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\left[\Gamma^{\lambda}{}_{k\mu}\frac{\partial x^{\rho\prime}}{\partial x^{\nu}}+\Gamma^{\lambda}{}_{k\nu}\frac{\partial x^{\rho\prime}}{\partial x^{\mu}}+\Gamma^{\lambda}{}_{\mu\nu}\frac{\partial x^{\rho\prime}}{\partial x^{k}}\right].$$

We now subtract from this expression the same expression with $k$ and $\nu$ interchanged

$$\begin{aligned}
&\frac{\partial^3 x^{\tau\prime}}{\partial x^k \partial x^\mu \partial x^\nu} - \frac{\partial^3 x^{\tau\prime}}{\partial x^\nu \partial x^\mu \partial x^k} = 0 = \\
&\frac{\partial x^{\tau\prime}}{\partial x^\lambda}\left[\left(\frac{\partial}{\partial x^k}\Gamma^\lambda{}_{\mu\nu}\right) + \Gamma^\lambda{}_{k\eta}\Gamma^\eta{}_{\mu\nu}\right] \\
&-\frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\frac{\partial x^{\eta\prime}}{\partial x^k}\left[\left(\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right) - \Gamma^{\tau\prime}{}_{\lambda\prime\sigma\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\rho\prime} - \Gamma^{\tau\prime}{}_{\rho\prime\lambda\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\sigma\prime}\right] \\
&-\frac{\partial x^{\sigma\prime}}{\partial x^\lambda}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\left[\Gamma^\lambda{}_{k\mu}\frac{\partial x^{\rho\prime}}{\partial x^\nu} + \Gamma^\lambda{}_{k\nu}\frac{\partial x^{\rho\prime}}{\partial x^\mu} + \Gamma^\lambda{}_{\mu\nu}\frac{\partial x^{\rho\prime}}{\partial x^k}\right] - \\
&\frac{\partial x^{\tau\prime}}{\partial x^\lambda}\left[\left(\frac{\partial}{\partial x^\nu}\Gamma^\lambda{}_{\mu k}\right) + \Gamma^\lambda{}_{\nu\eta}\Gamma^\eta{}_{\mu k}\right] \\
&+\frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^k}\frac{\partial x^{\eta\prime}}{\partial x^\nu}\left[\left(\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\right) - \Gamma^{\tau\prime}{}_{\lambda\prime\sigma\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\rho\prime} - \Gamma^{\tau\prime}{}_{\rho\prime\lambda\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\sigma\prime}\right] \\
&+\frac{\partial x^{\sigma\prime}}{\partial x^\lambda}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime}\left[\Gamma^\lambda{}_{\nu\mu}\frac{\partial x^{\rho\prime}}{\partial x^k} + \Gamma^\lambda{}_{\nu k}\frac{\partial x^{\rho\prime}}{\partial x^\mu} + \Gamma^\lambda{}_{\mu k}\frac{\partial x^{\rho\prime}}{\partial x^\nu}\right]
\end{aligned} \tag{7.9}$$

collecting all terms we find

$$\begin{aligned}
&\frac{\partial x^{\tau\prime}}{\partial x^\lambda}\left[\frac{\partial}{\partial x^k}\Gamma^\lambda{}_{\mu\nu} - \frac{\partial}{\partial x^\nu}\Gamma^\lambda{}_{\mu k} + \Gamma^\lambda{}_{k\eta}\Gamma^\eta{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\eta}\Gamma^\eta{}_{\mu k}\right] \\
&-\frac{\partial x^{\rho\prime}}{\partial x^\mu}\frac{\partial x^{\sigma\prime}}{\partial x^\nu}\frac{\partial x^{\eta\prime}}{\partial x^k}\left[\frac{\partial}{\partial x^{\eta\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\sigma\prime} - \frac{\partial}{\partial x^{\sigma\prime}}\Gamma^{\tau\prime}{}_{\rho\prime\eta\prime} + \Gamma^{\tau\prime}{}_{\lambda\prime\eta\prime}\Gamma^{\lambda\prime}{}_{\sigma\prime\rho\prime} - \Gamma^{\tau\prime}{}_{\lambda\prime\sigma\prime}\Gamma^{\lambda\prime}{}_{\eta\prime\rho\prime}\right] = 0.
\end{aligned} \tag{7.10}$$

If we now define the following [1]

$$R^\lambda{}_{\mu\nu k} = -\left[\frac{\partial}{\partial x^k}\Gamma^\lambda{}_{\mu\nu} - \frac{\partial}{\partial x^\nu}\Gamma^\lambda{}_{\mu k} + \Gamma^\lambda{}_{k\eta}\Gamma^\eta{}_{\mu\nu} - \Gamma^\lambda{}_{\nu\eta}\Gamma^\eta{}_{\mu k}\right], \tag{7.11}$$

[1] 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!

we can write eq. (7.10) as the transformation law for the tensor

$$R^{\sigma\prime}{}_{\alpha\prime\beta\prime\gamma\prime} = \frac{\partial x^{\sigma\prime}}{\partial x^{\lambda}} \frac{\partial x^{\mu}}{\partial x^{\alpha\prime}} \frac{\partial x^{\nu}}{\partial x^{\beta\prime}} \frac{\partial x^{k}}{\partial x^{\gamma\prime}} R^{\lambda}{}_{\mu\nu k}. \tag{7.12}$$

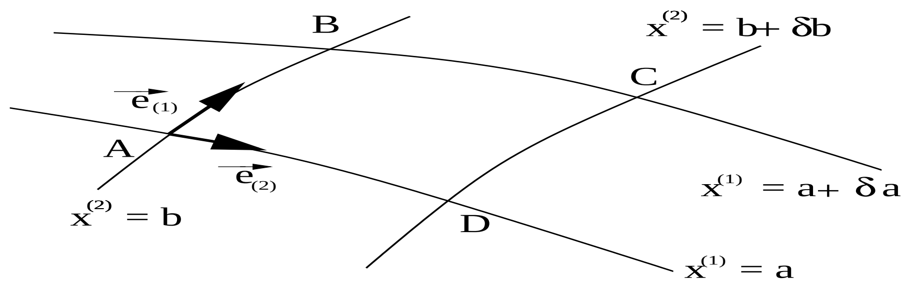
The tensor (7.11) is **The Curvature Tensor**, also called **The Riemann Tensor**, and it can be shown that it is the only tensor that can be constructed by using the metric, its first and second derivatives, and which is linear in the second derivatives.

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. (7.1) the part which does not transform as a tensor, and we do this in eq. (7.9).

## 7.2 b) The curvature tensor and the curvature of the spacetime

We shall now rederive the curvature tensor in a different way that explicitely shows why it espresses 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 coordinates lines $x^1 = a$, $x^1 = a + \delta a$, $x^2 = b$, $x^2 = b + \delta b$




Take a generic vector $\vec{V}$ and parallely transport $\vec{V}$ along AB, i.e. consider $\nabla_{\vec{e}_{(1)}} \vec{V} = 0$. From eq. (6.10) it follows that

$$e^{\beta}_{(1)} V^{\alpha}{}_{;\beta} = 0. \tag{7.13}$$

Since $\vec{e}_{(1)}$ has only $e^1_{(1)} \neq 0$ then

$$\frac{\partial V^\alpha}{\partial x^1} + \Gamma^{\alpha}{}_{\mu 1} V^\mu = 0. \tag{7.14}$$

This equation can be integrated along the line AB:

$$\delta V^\alpha_{AB} = - \int_{A(x^2=b)}^{B} \Gamma^{\alpha}{}_{\mu 1} V^\mu dx^1. \tag{7.15}$$

In a similar way, if we go from B to C along the line $x^1 = a + \delta a$

$$\frac{\partial V^\alpha}{\partial x^2} = - \Gamma^{\alpha}{}_{\mu 2} V^\mu \qquad \rightarrow \qquad \delta V^\alpha_{BC} = - \int_{B(x^1=a+\delta a)}^{C} \Gamma^{\alpha}{}_{\mu 2} V^\mu dx^2. \tag{7.16}$$

From C to D

$$\frac{\partial V^\alpha}{\partial x^1} = -\Gamma^\alpha{}_{\mu 1} V^\mu \qquad \rightarrow \qquad \delta V^\alpha_{CD} = -\int_{C(x^2=b+\delta b)}^{D} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1, \tag{7.17}$$

and from D back to A

$$\frac{\partial V^\alpha}{\partial x^2} = -\Gamma^\alpha{}_{\mu 2} V^\mu \qquad \rightarrow \qquad \delta V^\alpha_{DA} = -\int_{D(x^1=a)}^{A} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2. \tag{7.18}$$

The change in $\vec{V}$ due to this parallel transport will be a vector $\delta\vec{V}$ whose components can be found by addinge eqs. (7.15)-(7.18):

$$\begin{aligned} \delta V^\alpha = & -\int_{D(x^1=a)}^{A} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 \\ & -\int_{B(x^1=a+\delta a)}^{C} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 - \int_{C(x^2=b+\delta b)}^{D} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 \\ & -\int_{A(x^2=b)}^{B} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1. \end{aligned} \tag{7.19}$$

If the spacetime is flat $V^\mu$ does not change when paralleley transported, and $\delta V^\alpha = 0$. **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. (7.19) to first order in $\delta a$ and $\delta b$:

$$\begin{aligned} \delta V^\alpha \simeq & -\int_{D(x^1=a)}^{A} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 - \\ & \left[\int_{B(x^1=a)}^{C} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 + \frac{\partial}{\partial x^1}\left(\int_{B}^{C} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2\right)\delta a\right] \\ & -\left[\int_{C(x^2=b)}^{D} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 + \frac{\partial}{\partial x^2}\left(\int_{C}^{D} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1\right)\delta b\right] \\ & -\int_{A(x^2=b)}^{B} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1, \end{aligned} \tag{7.20}$$

Since

$$A=(a,b), \quad C=(a+\delta a, b+\delta b), \quad B=(a+\delta a, b), \quad \text{and} \quad D=(a, b+\delta b), \tag{7.21}$$

the previous equation becomes

$$\begin{aligned} \delta V^\alpha \simeq & + \int_b^{b+\delta b} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 \\ & - \int_b^{b+\delta b} \Gamma^\alpha{}_{\mu 2} V^\mu dx^2 - \left[ \int_b^{b+\delta b} \frac{\partial}{\partial x^1} \left( \Gamma^\alpha{}_{\mu 2} V^\mu \right) dx^2 \right] \delta a \\ & + \int_a^{a+\delta a} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1 + \left[ \int_a^{a+\delta a} \frac{\partial}{\partial x^2} \left( \Gamma^\alpha{}_{\mu 1} V^\mu \right) dx^1 \right] \delta b \\ & - \int_a^{a+\delta a} \Gamma^\alpha{}_{\mu 1} V^\mu dx^1, \end{aligned} \tag{7.22}$$

i.e.

$$\begin{aligned} \delta V^\alpha \simeq & -\delta a \int_b^{b+\delta b} \frac{\partial}{\partial x^1} \left( \Gamma^\alpha{}_{\mu 2} V^\mu \right) dx^2 \\ & + \delta b \int_a^{a+\delta a} \frac{\partial}{\partial x^2} \left( \Gamma^\alpha{}_{\mu 1} V^\mu \right) dx^1 \simeq \delta a \delta b \left[ -\frac{\partial}{\partial x^1} \left( \Gamma^\alpha{}_{\mu 2} V^\mu \right) + \frac{\partial}{\partial x^2} \left( \Gamma^\alpha{}_{\mu 1} V^\mu \right) \right]. \end{aligned} \tag{7.23}$$

Eq. (7.23) can be further developed by using eq. (7.14)

$$\frac{\partial V^\nu}{\partial x^1} = -\Gamma^\nu{}_{\mu 1} V^\mu, \qquad \frac{\partial V^\nu}{\partial x^2} = -\Gamma^\nu{}_{\mu 2} V^\mu \tag{7.24}$$

and it becomes

$$\begin{aligned} \delta V^\alpha &= \delta a \delta b \left[ \frac{\partial \Gamma^\alpha{}_{\mu 1}}{\partial x^2} V^\mu + \Gamma^\alpha{}_{\nu 1} \frac{\partial V^\nu}{\partial x^2} - \frac{\partial \Gamma^\alpha{}_{\mu 2}}{\partial x^1} V^\mu - \Gamma^\alpha{}_{\nu 2} \frac{\partial V^\nu}{\partial x^1} \right] \\ &= \delta a \delta b \left[ \frac{\partial \Gamma^\alpha{}_{\mu 1}}{\partial x^2} - \frac{\partial \Gamma^\alpha{}_{\mu 2}}{\partial x^1} - \Gamma^\alpha{}_{\nu 1} \Gamma^\nu{}_{\mu 2} + \Gamma^\alpha{}_{\nu 2} \Gamma^\nu{}_{\mu 1} \right] V^\mu. \end{aligned} \tag{7.25}$$

What does that mean? The term in parenthesis is a number, thus eq. (7.25) says that the $\delta V^\alpha$ are a linear combination of $V^\mu$. The indices 1 and 2 appear because we have choosen the path along $x^1$ and $x^2$ coordinates. **Note also that it is antisymmetric in 1 and 2** (if we interchange 1 and 2 $\delta V^\alpha$ reverse their sign, and this because we would go around the loop in opposite direction). If in place of $x^1$ and $x^2$, we choose arbitrary coordinate lines $x^\sigma$ and $x^\lambda$ we find

$$\delta V^\alpha = \delta a \delta b \left[ \frac{\partial \Gamma^\alpha{}_{\mu\sigma}}{\partial x^\lambda} - \frac{\partial \Gamma^\alpha{}_{\mu\lambda}}{\partial x^\sigma} - \Gamma^\alpha{}_{\nu\sigma}\Gamma^\nu{}_{\mu\lambda} + \Gamma^\alpha{}_{\nu\lambda}\Gamma^\nu{}_{\mu\sigma} \right] V^\mu. \tag{7.26}$$

The term in parenthesis coincides with the definition of the Riemann tensor (7.11). If we would go around the loop in the opposite direction the sign would reverse. This shows that the sign can be chosen arbitrarily, and this is the reason why the definition of the Riemann tensor given in textbooks may differ for a sign. In the following we shall assume that the curvature tensor is

$$R^\alpha{}_{\beta\mu\nu} = \Gamma^\alpha{}_{\beta\nu,\mu} - \Gamma^\alpha{}_{\beta\mu,\nu} + \Gamma^\alpha{}_{\sigma\mu}\Gamma^\sigma{}_{\beta\nu} - \Gamma^\alpha{}_{\sigma\nu}\Gamma^\sigma{}_{\beta\mu}. \tag{7.27}$$

We have already shown that this is a tensor by looking at the way it transforms under a coordinate transformation (eq. 7.12). But we want to see if it also agrees with the definition of tensors given in chapter 4. Let us contract eq. (7.26) with $V_\alpha$.

$$\delta V^\alpha V_\alpha = \delta a \delta b \left[ \frac{\partial \Gamma^\alpha{}_{\mu\sigma}}{\partial x^\lambda} - \frac{\partial \Gamma^\alpha{}_{\mu\lambda}}{\partial x^\sigma} - \Gamma^\alpha{}_{\nu\sigma}\Gamma^\nu{}_{\mu\lambda} + \Gamma^\alpha{}_{\nu\lambda}\Gamma^\nu{}_{\mu\sigma} \right] V^\mu V_\alpha. \tag{7.28}$$

The result of this contraction is, of course, a number. First we note that (7.28) is linear in $V^\alpha$, $V_\alpha, \delta a, \delta b$. $\delta a$ and $\delta b$ are the displacement along the basis vectors $\vec{e}_{(\sigma)}$ and $\vec{e}_{(\lambda)}$, thus $\delta V^\alpha$ depends linearly on $\delta a\vec{e}_{(\sigma)}$ and $\delta b\vec{e}_{(\lambda)}$. If we consider a generic $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ 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 $\vec{V}, \delta a\vec{e}_{(\sigma)}, \delta b\vec{e}_{(\lambda)}$), it will produce the following number

$$T(\tilde{V}, \vec{V}, \delta a\vec{e}_{(\sigma)}, \delta b\vec{e}_{(\lambda)}) = T^\alpha{}_{\beta\rho\delta} V_\alpha V^\beta \delta a e^\sigma_{(\rho)} \delta b e^\lambda_{(\delta)}. \tag{7.29}$$

Eq. (7.29) has the same structure of eq. (7.28). Therefore we are entitled to define the components of the Riemann tensor as in eq. (7.27).

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

$$\delta V^\alpha = 0 \qquad \to \qquad R^\alpha{}_{\beta\gamma\delta} = 0, \tag{7.30}$$

**in any reference frame**. The components of the Riemann tensor assume a very nice form when computed in a locally inertial frame:

$$R^\alpha{}_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\sigma} \left[ g_{\sigma\nu,\beta\mu} - g_{\sigma\mu,\beta\nu} + g_{\beta\mu,\sigma\nu} - g_{\beta\nu,\sigma\mu} \right], \tag{7.31}$$

or lowering the index $\alpha$

$$R_{\alpha\beta\mu\nu} = g_{\alpha\lambda} R^\lambda{}_{\beta\mu\nu} = \frac{1}{2} \left[ g_{\alpha\nu,\beta\mu} - g_{\alpha\mu,\beta\nu} + g_{\beta\mu,\alpha\nu} - g_{\beta\nu,\alpha\mu} \right]. \tag{7.32}$$

It should be stressed that

1) The Riemann tensor is linear in the second derivatives of $g_{\mu\nu}$.

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.

## 7.3 Symmetries

From eq. (7.32) it is easy to verify that

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta}, \tag{7.33}$$

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0. \tag{7.34}$$

Since $R_{\alpha\beta\mu\nu}$ is a tensor, these symmetry properties are valid in any references frame. The symmetries of the Riemann tensor reduce the number of independent components to 20.

## 7.4 The Riemann tensor gives the commutator of covariant derivatives

Let us consider the second covariant derivatives of a vector field $\vec{V}$

$$\nabla_\alpha \nabla_\beta V^\mu = \nabla_\alpha (V^\mu{}_{;\beta}) = (V^\mu{}_{;\beta})_{,\alpha} + \Gamma^\mu{}_{\sigma\alpha} V^\sigma{}_{;\beta} - \Gamma^\sigma{}_{\beta\alpha} V^\mu{}_{;\sigma}. \tag{7.35}$$

In a locally inertial frame $\Gamma^\mu{}_{\sigma\alpha} = 0$, and eq. (7.35) becomes

$$\nabla_\alpha \nabla_\beta V^\mu = (V^\mu{}_{;\beta})_{,\alpha} = V^\mu{}_{,\beta,\alpha} + \Gamma^\mu{}_{\nu\beta,\alpha} V^\nu. \tag{7.36}$$

By interchanging $\alpha$ and $\beta$

$$\nabla_\beta \nabla_\alpha V^\mu = (V^\mu{}_{;\alpha})_{,\beta} = V^\mu{}_{,\alpha,\beta} + \Gamma^\mu{}_{\nu\alpha,\beta} V^\nu. \tag{7.37}$$

The commutator of the covariant derivatives then is

$$[\nabla_\alpha, \nabla_\beta]\, V^\mu = \nabla_\alpha \nabla_\beta V^\mu - \nabla_\beta \nabla_\alpha V^\mu = (\Gamma^\mu{}_{\nu\beta,\alpha} - \Gamma^\mu{}_{\nu\alpha,\beta})\, V^\nu. \tag{7.38}$$

Since in a locally inertial frame

$$R^\lambda{}_{\mu\nu k} = \Gamma^\lambda{}_{\mu k,\nu} - \Gamma^\lambda{}_{\mu\nu,k} \tag{7.39}$$

(equivalent to eq. 7.32), eq. (7.38) becomes

$$[\nabla_\alpha, \nabla_\beta]\, V^\mu = R^\mu{}_{\nu\alpha\beta} V^\nu. \tag{7.40}$$

This is a tensor equation and since it is valid in a given reference frame, it will be valid in **any** frame. Eq. (7.40) implies that in curved spacetime covariant derivatives **do not commute** and therefore the order in which they appear is important.

## 7.5 The Bianchi identities

Let us differentiate eq. (7.32) with respect to $x^\lambda$ (and rememeber that it is valid in a locally inertial frame)

$$R_{\alpha\beta\mu\nu,\lambda} = \frac{1}{2}\left[g_{\alpha\nu,\beta\mu\lambda} - g_{\alpha\mu,\beta\nu\lambda} + g_{\beta\mu,\alpha\nu\lambda} - g_{\beta\nu,\alpha\mu\lambda}\right]. \tag{7.41}$$

By using the fact that $g_{\alpha\beta}$ is symmetric and eq. (7.41) one can show that

$$R_{\alpha\beta\mu\nu,\lambda} + R_{\alpha\beta\lambda\mu,\nu} + R_{\alpha\beta\nu\lambda,\mu} = 0. \tag{7.42}$$

Since it is valid in a locally inertial frame and it is a tensor equation, it will be valid in any frame:

$$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0, \tag{7.43}$$

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 8

# 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 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 what to put on the right-hand-side of the equations as a source of the gravitational field.

## 8.1 The stress-energy tensor in Special Relativity

In Special Relativity, we define the **energy-momentum** four-vector of a particle of mass $m$ and velocity $\vec{v} = \frac{d\vec{x}}{dt}$ in the following way

$$p^\alpha = mcu^\alpha, \qquad \alpha = 0, 3, \tag{8.1}$$

where $u^\alpha = \frac{dx^\alpha}{d\tau}$ is the four-velocity, $\tau$ is the proper time

$$\begin{aligned} d\tau^2 &= -\eta_{\alpha\beta} dx^\alpha dx^\beta, \\ d\tau &= (c^2 dt^2 - dx^2 - dy^2 - dz^2)^{1/2} = (1 - \frac{v^2}{c^2})^{1/2} c dt, \end{aligned} \tag{8.2}$$

and

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{8.3}$$

The time-component of the energy-momentum vector does represent the energy of the particle

$$p^0 = \frac{E}{c}, \qquad \gamma = (1 - v^2/c^2)^{-1/2}, \quad \text{and} \quad E = mc^2\gamma. \tag{8.4}$$

The space-components are the components of the three-dimensional momentum

$$\vec{p} = m\gamma\vec{v}. \tag{8.5}$$

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 $x_n(t)$, each with an energy-momentum vector $p_n^\alpha$. (We will define the stress-energy tensor for fluids and for electromagnetic fields when we will study stars and charged black holes).

We define the **density of energy** as

$$T^{00} \equiv \sum_n c p_n^0(t) \delta^3(\mathbf{x} - \mathbf{x}_n(t)) = \sum_n E_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \tag{8.6}$$

the **density of momentum** $\frac{1}{c} T^{0i}$, where $T^{0i}$ is defined as

$$T^{0i} \equiv \sum_n c p_n^i(t) \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \qquad i = 1, 3 \tag{8.7}$$

and the **current** of momentum as

$$T^{ki} \equiv \sum_n p_n^k(t) \frac{dx_n^i(t)}{dt} \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \qquad k = 1, 3 \quad i = 1, 3. \tag{8.8}$$

$\delta^3(\mathbf{x} - \mathbf{x}_n)$ is the Dirac delta-function defined by the statement that for any smooth function $f(x)$

$$\int_{-\infty}^{\infty} d^3x f(x) \delta^3(\mathbf{x} - \mathbf{y}) = f(\mathbf{y}), \tag{8.9}$$

where, if $\mathbf{y} = (x0, y0, z0)$

$$\delta^3(\mathbf{x} - \mathbf{y}) = \delta(x - x0)\delta(y - y0)\delta(z - z0), \tag{8.10}$$

or, in polar coordinates

$$\delta^3(\mathbf{r} - \mathbf{r0}) = \frac{1}{r^2 \sin^2\theta}\delta(r - r0)\delta(\theta - \theta_0)\delta(\varphi - \varphi_0). \tag{8.11}$$

It should be noted that according to the definition (8.9) the three-dimensional $\delta$-function has the dimensions of the inverse of a cubic lenght $l^{-3}$. For this reason, for example, since $T^{00}$ has the dimensions of an energy per unit volume, it is the *density* of energy. [1] The two definitions (8.7) and (8.8) can be unified into a single formula

$$T^{\alpha\beta} = \sum_n p_n^\alpha \frac{dx_n^\beta(t)}{dt}\delta^3(\mathbf{x} - \mathbf{x}_n(t)), \qquad \alpha, \beta = 0, 3. \tag{8.12}$$

Since

$$p_n^\alpha = \frac{E_n}{c^2}\frac{dx_n^\alpha(t)}{dt}, \tag{8.13}$$

eq. (8.12) can also be written as

[1] Properties of the $\delta$-function

$$\delta(x) = \delta(-x), \qquad \delta(cx) = \frac{1}{|c|}\delta(x)$$

$$\delta[g(x)] = \sum_j \frac{1}{|g\prime(x^j)|}\delta(x - x^j) \qquad x\delta(x) = 0$$

$$\int_{-\infty}^{\infty} d^3x f(x)\delta'(\mathbf{x} - \mathbf{y}) = -f'(y).$$

$$T^{\alpha\beta} = c^2 \sum_n \frac{p_n^\alpha p_n^\beta}{E_n} \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \tag{8.14}$$

which clearly shows that $T^{\alpha\beta}$ is **symmetric**

$$T^{\alpha\beta} = T^{\beta\alpha}. \tag{8.15}$$

Finally, an alternative way of writing eq. (8.12) can be found by using the property (8.9) of the $\delta$-function

$$T^{\alpha\beta} = c \sum_n \int p_n^\alpha \frac{dx_n^\beta}{d\tau_n} \delta^4(\mathbf{x} - \mathbf{x}_n(\tau_n)) d\tau_n, \tag{8.16}$$

where

$$\delta^4(\mathbf{x} - \mathbf{y}) = \delta(\tau_n - ct))\delta(x - x0)\delta(y - y0)\delta(z - z0), \tag{8.17}$$

The meaning of the different components is the following

$T^{00}$ = energy-density. In the non-relativistic case $v << c$, $p_n^0 \sim m_n c^2$ and $T^{00} \sim \sum_n m_n c^2 \delta^3(\mathbf{x} - \mathbf{x}_n(t))$ reduces to the density of matter $\rho c^2$ where

$$\rho = \sum_n m_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \tag{8.18}$$

(remember the dimensions of the $\delta$-function) .

$T^{0i}$ = energy which flows across the unit surface orthogonal to the axis $x^i$ per unit time (i=1,3) (see eq. (8.16)).

$T^{ik}$ = flux of the i- component of the three-momentum $\vec{p}$ across the unit surface orthogonal to the axis $x^k$ (i,k=1,3).

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) $T^{\alpha\beta}$ **is a tensor**

Let us subject the coordinates $x^\alpha$ to a Lorentz transformation

$$x^\alpha = L^\alpha{}_{\gamma'} x^{\gamma'}, \tag{8.19}$$

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

$$L^0{}_{0'} = \gamma, \quad L^0{}_{j'} = L^j_{0'} = \frac{\gamma}{c} v_j, \qquad L^i{}_{j'} = \delta^i{}_{j'} + \frac{\gamma - 1}{v^2} v_i v_j. \quad i, j = 1, 3 \tag{8.20}$$

and $v^i$ are the components of the velocity of the boost. Accordingly, the four-momentum will transform as

$$p_n^\alpha = L^\alpha{}_{\gamma'} p_n^{\gamma'}, \tag{8.21}$$

and $T^{\alpha\beta}$ becomes

$$\begin{aligned} T^{\alpha\beta} &= c \sum_n \int d\tau_n L^\alpha{}_{\gamma'} L^\beta{}_{\delta'} p_n^{\gamma'} \frac{dx_n^{\delta'}}{d\tau_n} \delta^4(\mathbf{x}' - \mathbf{x}_n(\tau_n)') d\tau_n = \\ &= L^\alpha{}_{\gamma'} L^\beta{}_{\delta'} c \sum_n \int d\tau_n p_n^{\gamma'} \frac{dx_n^{\delta'}}{d\tau_n} \delta^4(\mathbf{x}' - \mathbf{x}_n(\tau_n)') d\tau_n. \end{aligned} \tag{8.22}$$

Consequently $T^{\alpha\beta}$ is a tensor because it transforms as a tensor

$$T^{\alpha\beta} = L^\alpha{}_{\gamma'} L^\beta{}_{\delta'} T^{\gamma'\delta'}. \tag{8.23}$$

2) $T^{\alpha\beta}$ **does satisfy a conservation law**

In order to show it we differentiate eq. (8.8):

$$\frac{\partial T^{\alpha i}}{\partial x^i} = \sum_n p_n^\alpha(t) \frac{dx_n^i(t)}{dt} \frac{\partial}{\partial x^i} \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \tag{8.24}$$

where $\alpha = 0, 3$ and $i = 1, 3$. Since

$$\frac{\partial}{\partial x^i} \delta^3(\mathbf{x} - \mathbf{x}_n(t)) = -\frac{\partial}{\partial x_n^i} \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \tag{8.25}$$

eq. (8.24) becomes

$$\begin{aligned} \frac{\partial T^{\alpha i}}{\partial x^i} &= -\sum_n p_n^\alpha(t) \frac{dx_n^i(t)}{dt} \frac{\partial}{\partial x_n^i} \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \\ &= -\sum_n p_n^\alpha(t) \frac{dx_n^i(t)}{dt} \frac{\partial}{\partial t} \frac{\partial t}{\partial x_n^i} \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \\ &= -\sum_n p_n^\alpha(t) \frac{\partial}{\partial t} \delta^3(\mathbf{x} - \mathbf{x}_n(t)). \end{aligned} \tag{8.26}$$

By making use of eqs. (8.6) and (8.7), eq. (8.26) gives

$$\frac{\partial T^{\alpha i}}{\partial x^i} = -\frac{1}{c} \frac{\partial}{\partial t} T^{\alpha 0} + \sum_n \frac{dp_n^\alpha(t)}{dt} \delta^3(\mathbf{x} - \mathbf{x}_n(t)). \tag{8.27}$$

Since

$$\frac{dp_n^\alpha(t)}{dt} = \frac{dp_n^\alpha(\tau)}{d\tau} \frac{d\tau}{dt} = \frac{d\tau}{dt} f_n^\alpha, \tag{8.28}$$

where $f_n^\alpha$ is the relativistic force, the last term in eq. (8.27) can be considered as a *density of force* $G^\alpha$ defined as

$$G^\alpha(\mathbf{x}, t) = \sum_n \frac{dp_n^\alpha(t)}{dt} \delta^3(\mathbf{x} - \mathbf{x}_n(t)) = \sum_n \delta^3(\mathbf{x} - \mathbf{x}_n(t)) \frac{d\tau}{dt} f_n^\alpha. \tag{8.29}$$

It is a density because the $\delta$-function is $[l^{-3}]$. If the particles are free, $f_n^\alpha = 0$ and eq. (8.27) becomes

$$\frac{\partial}{\partial x^\beta} T^{i\beta} + \frac{1}{c}\frac{\partial}{\partial t} T^{0\beta} = \frac{\partial}{\partial x^\beta} T^{\alpha\beta} = 0, \tag{8.30}$$

or

$$T^{\alpha\beta}{}_{,\beta} = 0, \tag{8.31}$$

which is the conservation law we were looking for.

Why is $T^{\alpha\beta}{}_{,\beta} = 0$ a conservation law? To answer this question, let us start with a familiar equation in classical electrodynamics. Consider a collection of charged particles of density $\rho$ enclosed in a volume $V$.

$$\frac{\partial}{\partial t} \int_V \rho dV \tag{8.32}$$

will be the variation of charge inside the volume $V$. Be $S$ the surface enclosing the volume, and $\vec{n}$ the normal vector, which is assumed to be positive if pointing outward.

$$\rho \vec{v} \cdot \vec{n} dS \tag{8.33}$$

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

$$\int_S \rho \vec{v} \cdot \vec{n} dS \tag{8.34}$$

is the total charge per unit time, which flows across the surface $S$ enclosing the volume $V$. The continuity equation then says that

$$\frac{\partial}{\partial t} \int_V \rho dV = - \int_S \rho \vec{v} \cdot \vec{n} dS. \tag{8.35}$$

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

$$\vec{J} = \rho\vec{v}, \tag{8.36}$$

eq. (8.35) becomes

$$\frac{\partial}{\partial t}\int_V \rho dV = -\int_S \vec{J}\cdot\vec{n}dS. \tag{8.37}$$

We now apply the Gauss theorem:

$$\int_S \vec{J}\cdot\vec{n}dS = \int_V div\vec{J}dV, \tag{8.38}$$

and eq. (8.37) becomes

$$\frac{\partial}{\partial t}\int_V \rho dV = -\int_V div\vec{J}dV. \tag{8.39}$$

Since the volume $V$ is arbitrary, we can write

$$div\vec{J} = -\frac{\partial\rho}{\partial t}, \tag{8.40}$$

or

$$\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = -\frac{\partial\rho}{\partial t}, \tag{8.41}$$

which is the **continuity equation** in a differential form. Let us now transform eq. (8.40) in a four-dimensional form. We define a four-current

$$J^\alpha = \rho\frac{dx^\alpha}{dt} = (\rho c, \vec{J}), \tag{8.42}$$

Then eq. (8.40) becomes

$$\frac{\partial}{\partial x^\alpha}J^\alpha = 0, \qquad\qquad \alpha = 0, 3. \tag{8.43}$$

We are now going to show that any current $J^\alpha(x)$ which satisfies the conservation law (8.43) is associated to a total charge $Q$ defined as

$$Q = \int_V J^0 dV, \tag{8.44}$$

which is conserved. **The integral in eq. (8.44) is evaluated at some fixed time, thus we say that the integration is performed on an hypersurface $x^0 = const$ over the whole three-dimensional space.** The total charge $Q$ is a conserved quantity for the following reason. By virtue of eq. (8.43)

$$\frac{1}{c}\frac{dQ}{dt} = \int_{allspace} \frac{1}{c}\frac{\partial}{\partial t} J^0 dV = -\int_{allspace} div\vec{J} dV = -\int_{surface} J^k dS_k. \tag{8.45}$$

The last equality follows from the application of the Gauss theorem, and the subscript 'surface' means that we are considering the flux of $\vec{J}$ across the surface which encloses the whole space. $dS_k$ are the element of surface orthogonal to $x^k$. If $\vec{J}$ goes to zero at infinity, the last term in eq. (8.45) vanishes, and therefore the total charge $Q$ is a conserved quantity. It can also be shown that $Q$ is a scalar (see Weinberg pg. 41).

And now let us go back to equation (8.31). Let us assume for example that $\alpha = 0$:

$$\frac{\partial T^{0x}}{\partial x} + \frac{\partial T^{0y}}{\partial y} + \frac{\partial T^{0z}}{\partial z} = -\frac{1}{c}\frac{\partial T^{00}}{\partial t}. \tag{8.46}$$

If we integrate over a volume $V$ as we did before, we get

$$-\frac{1}{c}\frac{\partial}{\partial t}\int_V T^{00} dV = \int_V div(T^{0k}) dV = \int_S T^{0k} dS_k. \tag{8.47}$$

Remembering that $T^{00}$ is the energy-density and $T^{0k}$ is the energy which flows across the unit surface orthogonal to $x^k$ it is clear that eq. (8.47) 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. (8.44) we can define a vector

$$P^\alpha = \int_V T^{\alpha 0} dV, \qquad \alpha = 0, 3, \tag{8.48}$$

which can be identified as the conserved energy-momentum vector of the system. For example

$$P^0 = \int_V T^{00} dV, \tag{8.49}$$

does represent the total energy of the system. It is conserved because

$$\frac{1}{c}\frac{dP^0}{dt} = \frac{1}{c}\int_{allspace} \frac{\partial}{\partial t} T^{00} dV = -\int_{allspace} \frac{\partial}{\partial x^i} T^{0i} dV = -\int_{surface} T^{0i} dS_i = 0. \tag{8.50}$$

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:

## 8.2 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* $g_{\mu\nu} \to \eta_{\mu\nu}$ *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. (8.31) 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. (8.31) in the alternative form

$$T^{\alpha\beta}{}_{;\beta} = 0. \tag{8.51}$$

Then we observe that in the light of the Principle of General Covariance, since the conservation law (8.31) 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. (8.51).

**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

$$\Gamma^{\mu}{}_{\lambda\mu} = \frac{1}{2} g^{\mu m} \left( \frac{\partial g_{m\lambda}}{\partial x^{\mu}} + \frac{\partial g_{m\mu}}{\partial x^{\lambda}} - \frac{\partial g_{\lambda\mu}}{\partial x^{m}} \right). \tag{8.52}$$

The first and the third term give

$$g^{\mu m} \frac{\partial g_{m\lambda}}{\partial x^{\mu}} - g^{\mu m} \frac{\partial g_{\lambda\mu}}{\partial x^{m}} = g^{\mu m} \frac{\partial g_{m\lambda}}{\partial x^{\mu}} - g^{m\mu} \frac{\partial g_{\mu\lambda}}{\partial x^{m}} = 0, \tag{8.53}$$

due to the symmetry of $g_{\alpha\beta}$, therefore

$$\Gamma^{\mu}{}_{\lambda\mu} = \frac{1}{2} g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^{\lambda}}. \tag{8.54}$$

For any arbitrary matrix $M$

$$\mathrm{Tr}\left[ M^{-1}(x) \frac{\partial}{\partial x^{\lambda}} M(x) \right] = \frac{\partial}{\partial x^{\lambda}} ln[Det M(x)]. \tag{8.55}$$

But this is what we have on the right-hand side of eq. (8.54), therefore, if we call $Det(g) = -g$, eq. (8.54) becomes

$$\Gamma^{\mu}{}_{\lambda\mu} = \frac{1}{2} \frac{\partial}{\partial x^{\lambda}} ln[-g] = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^{\lambda}} \sqrt{-g}. \tag{8.56}$$

Thus for example, if $V^\mu$ is a vector

$$V^\lambda{}_{;\lambda} = V^\lambda{}_{,\lambda} + \Gamma^\lambda{}_{\alpha\lambda}V^\alpha = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\lambda}\left(\sqrt{-g}V^\lambda\right), \tag{8.57}$$

and for $T^{\mu\nu}$

$$T^{\mu\nu}{}_{;\mu} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\lambda}(\sqrt{-g}T^{\mu\nu}) + \Gamma^\nu{}_{\lambda\mu}T^{\mu\lambda}. \tag{8.58}$$

In particular, if $F^{\mu\nu}$ is antisymmetric, the last term in eq. (8.58) is zero and

$$F^{\mu\nu}{}_{;\mu} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^\lambda}(\sqrt{-g}F^{\mu\nu}). \tag{8.59}$$

Now we go back to eq. (8.51). By using eq. (8.58) it becomes

$$\frac{\partial}{\partial x^\lambda}(\sqrt{-g}T^{\mu\nu}) = -\sqrt{-g}\Gamma^\nu{}_{\lambda\mu}T^{\mu\lambda}, \tag{8.60}$$

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

$$P^\alpha = \int_V \sqrt{-g}T^{\alpha 0}dV, \qquad \alpha = 0,3, \tag{8.61}$$

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 9

# 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 **is not** a source for the electromagnetic field, and it does not appear on the right-hand side of the equations. In gravity the situation is different. The equation

$$E = mc^2, \tag{9.1}$$

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 masses arbitrarily moving 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

$$\nabla^2\Phi = 4\pi G\rho, \tag{9.2}$$

where $\rho$ is the matter density, $\Phi$ is the newtonian potential and $\nabla^2$ is the Laplace operator in cartesian coordinates

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x^2}. \tag{9.3}$$

Let us start by asking how the equations should look in the weak field limit.

## 9.1 The geodesic equations in the weak field limit

Consider a non-relativistic particle which moves in a **weak and stationary** gravitational field. Be $\tau$ the proper time. Since $v << c$ , it follows that

$$\frac{dx^i}{dt} << c \qquad \rightarrow \qquad \frac{dx^i}{d\tau} << \frac{cdt}{d\tau} = \frac{dx^0}{d\tau}. \tag{9.4}$$

In an arbitrary coordinate system the geodesic equations are

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \quad \rightarrow \quad \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{cdt}{d\tau}\right)^2 = 0. \tag{9.5}$$

From the expressions of the affine connections in terms of $g_{\mu\nu}$ we easily find that

$$\Gamma^\mu_{00} = \frac{1}{2}g^{\mu\sigma}\left(2g_{0\sigma,0} - g_{00,\sigma}\right). \tag{9.6}$$

In addition, if the field is stationary $g_{0\sigma,0} = 0$ , and

$$\Gamma^\mu_{00} = -\frac{1}{2}g^{\mu\sigma}\frac{\partial g_{00}}{\partial x^\sigma}. \tag{9.7}$$

Since we have assumed that the gravitational field is weak, we can choose a coordinate system such that

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| << 1, \tag{9.8}$$

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. (9.8) into eq. (9.7), and retain only the terms up to first order in $h_{\mu\nu}$ we find

$$\Gamma^\mu_{00} \sim -\frac{1}{2}\eta^{\mu\sigma}\frac{\partial h_{00}}{\partial x^\sigma}, \tag{9.9}$$

and the geodesic equation becomes

$$\frac{d^2x^\mu}{d\tau^2} = \frac{1}{2}\eta^{\mu\alpha}\frac{\partial h_{00}}{\partial x^\alpha}\left(\frac{cdt}{d\tau}\right)^2, \tag{9.10}$$

or, splitting the time- and the space-components

$$\frac{d^2\vec{x}}{d\tau^2} = \frac{1}{2}\vec{\nabla}h_{00}\left(\frac{cdt}{d\tau}\right)^2, \qquad \text{and} \qquad \frac{d^2ct}{d\tau^2} = -\frac{1}{2}\frac{\partial h_{00}}{\partial ct}\left(\frac{cdt}{d\tau}\right)^2 = 0, \tag{9.11}$$

where

$$\vec{\nabla} \to \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\frac{\partial}{\partial z}\right) \tag{9.12}$$

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 $\frac{cdt}{d\tau} = 1$ and the first of eqs. (9.11) becomes

$$\frac{d^2\vec{x}}{d\tau^2} = \frac{1}{2}\vec{\nabla}h_{00}. \tag{9.13}$$

We should remember that the corresponding newtonian equation is

$$\frac{d^2\vec{x}}{dt^2} = -\vec{\nabla}\Phi, \tag{9.14}$$

where $\Phi$ is the gravitational potential given by the Poisson equation (9.2). By comparing eqs. (9.14) and (9.13), and since $\tau = ct$ we see that it must be

$$h_{00} = -2\frac{\Phi}{c^2} + const. \tag{9.15}$$

For example if the field is stationary and spherically symmetric, the newtonian potential is

$$\Phi = -\frac{GM}{r}, \tag{9.16}$$

and if we require that $h_{00}$ vanishes at infinity, the constant must be zero and eq. (9.15) gives

$$h_{00} = -2\frac{\Phi}{c^2}, \qquad \text{and} \qquad g_{00} = -(1 + 2\frac{\Phi}{c^2}). \tag{9.17}$$

**Thus we have shown that in the weak field limit the geodesic equations reduce to the newtonian law of gravitation.** This suggest 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 (9.2) could be

$$\nabla^2 g_{00} = -\frac{8\pi G}{c^4} T_{00}. \tag{9.18}$$

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

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{9.19}$$

where $G_{\mu\nu}$ is an operator which acts on $g_{\mu\nu}$ which we shall now define. It should be stressed that, by the Principle of General Covariance, if equation (9.19) holds in a given reference frame, it will hold in any other frame.

## 9.2 The 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

$$\frac{\partial^3 g_{\mu\nu}}{\partial x_\mu^3}, \qquad \frac{\partial^2 g_{\mu\nu}}{\partial x_\mu^2} \cdot \frac{\partial g_{\mu\nu}}{\partial x_\nu}, \qquad \frac{\partial g_{\mu\nu}}{\partial x_\nu}, \tag{9.20}$$

then, in order to be dimensionally homogeneous each term should be multiplied by a constant having the dimensions of a suitable power of a lenght

$$\frac{\partial^3 g_{\mu\nu}}{\partial x_\mu^3} \cdot l, \qquad \frac{\partial^2 g_{\mu\nu}}{\partial x_\mu^2} \frac{\partial g_{\mu\nu}}{\partial x_\nu} \cdot l, \qquad \frac{\partial g_{\mu\nu}}{\partial x_\nu} \cdot \frac{1}{l}. \tag{9.21}$$

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.

$$G^{\mu\nu}{}_{;\nu} = 0. \tag{9.22}$$

5) In the weak field limit it must reduce to (compare with eq. (9.18)

$$G_{00} \sim -\nabla^2 g_{00}. \tag{9.23}$$

In this last assumption the Principle of Equivalence and the weak field limit explicitely appear.

In the preceeding 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. (7.11), and it contains the information on the gravitational field. However we cannot use it directely in the field equations we are looking for, since it has four indices (it is a $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ tensor) while we need a $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ (or $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$) tensor. In addition, it does not satisfy the same conservation law as the stress-energy tensor (8.30). However it can be shown that there is only one $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ tensor, and only one scalar which can be constructed starting from the Riemann tensor and contracting it with the metric:

**the Ricci tensor**:

$$R_{\mu\nu} = g^{k\alpha} R_{k\mu\alpha\nu} = R^{\alpha}{}_{\mu\alpha\nu}, \tag{9.24}$$

which is a symmetric tensor because of the symmetry property of the Riemann tensor

$$R_{k\mu\alpha\nu} = R_{\alpha\nu k\mu}, \tag{9.25}$$

and the **scalar curvature**

$$R = R^{\alpha}{}_{\alpha}. \tag{9.26}$$

(remember that the contraction in eq. (9.26) has the following meaning

$$R^{\alpha}{}_{\alpha} = R^{0}{}_{0} + R^{1}{}_{1} + R^{2}{}_{2} + R^{3}{}_{3}). \tag{9.27}$$

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 $R_{k\mu\alpha\nu}$ with the metric (see Weinberg pg. 143). 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

$$G_{\mu\nu} = C_1 R_{\mu\nu} + C_2 g_{\mu\nu} R, \tag{9.28}$$

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

$$G^{\mu\nu}{}_{;\mu} = C_1 R^{\mu\nu}{}_{;\mu} + C_2 g^{\mu\nu} R_{;\mu} = 0. \tag{9.29}$$

(remember that the covariant derivative of $g_{\mu\nu}$ vanishes). Now a very remarkable thing happens: eq. (9.29) is satisfied because of the Bianchi identities

$$R_{\lambda\mu\nu k;\eta} + R_{\lambda\mu\eta\nu;k} + R_{\lambda\mu k\eta;\nu} = 0. \tag{9.30}$$

In fact by contracting these equations we find

$$g^{\lambda\nu}\left(R_{\lambda\mu\nu k;\eta} + R_{\lambda\mu\eta\nu;k} + R_{\lambda\mu k\eta;\nu}\right) = g^{\lambda\nu}\left(R_{\lambda\mu\nu k;\eta} - R_{\lambda\mu\nu\eta;k} + R_{\lambda\mu k\eta;\nu}\right) \tag{9.31}$$

$$\left(R_{\mu k;\eta} - R_{\mu\eta;k} + R^{\nu}{}_{\mu k\eta;\nu}\right) = 0.$$

Contracting again

$$g^{\mu k}\left(R_{\mu k;\eta} - R_{\mu\eta;k} + R^{\nu}{}_{\mu k\eta;\nu}\right) = R_{;\eta} - R^{k}{}_{\eta;k} - R^{\nu}{}_{\eta;\nu} = 0. \tag{9.32}$$

The last expression can be rewritten in the following form

$$\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)_{;\nu} = 0. \tag{9.33}$$

Therefore, the Bianchi identities say that if

$$\frac{C_2}{C_1} = -\frac{1}{2}, \tag{9.34}$$

eq. (9.33) will be satisfied. We still need $C_1$.

In the weak field limit [1]

$$|T_{ij}| << |T_{00}|, \qquad i,j=1,3, \tag{9.37}$$

and therefore

$$|G_{ij}| << |G_{00}|, \qquad i,j=1,3. \tag{9.38}$$

From eqs. (9.28) and (9.34) it follows

$$|C_1\left(R_{ij}-\frac{1}{2}g_{ij}R\right)| << |G_{00}|, \tag{9.39}$$

hence

$$R_{ij} \simeq \frac{1}{2}g_{ij}R. \tag{9.40}$$

Since $g_{ij} \simeq \eta_{ij}$

$$R_{kk} \simeq \frac{1}{2}R, \qquad\qquad k=1,3 \tag{9.41}$$

consequently

$$R = g^{\mu\nu}R_{\mu\nu} \simeq \eta^{\mu\nu}R_{\mu\nu} = -R_{00} + \sum_k R_{kk} = -R_{00} + \frac{3}{2}R, \tag{9.42}$$

[1] The fact that in the weak field limit $T_{\mu\nu} << T_{00}$ can be easily understood if we consider a system on non-interacting particles. If $\rho$ is the mass density

$$\rho = \sum_i m_i \delta(\vec{r}-\vec{r}_i), \tag{9.35}$$

where $\vec{r}_i$ denotes the positions of the particles, the stress-energy tensor (8.12) can be also written as

$$T^{\mu\nu} = \rho c^2 \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}. \tag{9.36}$$

It is clear that, if $\frac{dx^i}{d\tau} << \frac{dx^0}{d\tau}$ $i=1,3$ the dominant term will be $T^{00}$ .

and

$$R \simeq 2R_{00}. \tag{9.43}$$

Since

$$G_{\mu\nu} = C_1 \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right), \tag{9.44}$$

we find

$$G_{00} \simeq C_1 2 R_{00}. \tag{9.45}$$

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 we get

$$R_{00} \simeq -\frac{1}{2} \frac{\partial^2 g_{00}}{\partial x^i \partial^j} = -\frac{1}{2} \nabla^2 g_{00}, \qquad i, k = 1, 3 \tag{9.46}$$

namely

$$G_{00} \simeq -C_1 \nabla^2 g_{00}, \tag{9.47}$$

A comparison of this equation with eq. (9.23) shows that if we require that the relativistic field equations reduce to the newtonian equations in the weak field limit it must be

$$C_1 = 1. \tag{9.48}$$

In conclusion, the Einstein's field equations are [2]

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{9.49}$$

[2] Although we call these equations the Einstein equations, they were derived independently (and in a more elegant form) by D. Hilbert in the same year. However Einstein showed the implications of these equations in the theory of the solar system, and in par-

where

$$G_{\mu\nu} = \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right), \tag{9.50}$$

and it is called **The Einstein tensor**. An alternative form is

$$R_{\mu\nu} = \frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right). \tag{9.51}$$

In vacuum $T_{\mu\nu} = 0$ and the Einstein equations reduce to

$$R_{\mu\nu} = 0. \tag{9.52}$$

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. (9.49) the following term

$$\left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - \lambda g_{\mu\nu} \right) = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{9.53}$$

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.

ticular that the precession of the perihelion of Mercury has a relativistic origin. This led to the theory's acceptance and since then the equations have been called the Einstein equations.

## 9.3 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 $x^{\mu\prime} = x^{\mu\prime}(x^\alpha)$ the 'transformed' tensor $g'_{\mu\nu} = g_{\mu\prime\nu\prime}$ 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 $x^{\mu\prime}(x^\alpha)$, 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

$$\Box A_\alpha - \frac{\partial^2 A^\beta}{\partial x^\alpha \partial x^\beta} = -\frac{4\pi}{c} J_\alpha. \tag{9.54}$$

(where $\Box = -\frac{\partial^2}{c^2 \partial t^2} + \nabla^2 = \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta}$ ). 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

$$J^\alpha{}_{,\alpha} = 0, \qquad \text{i.e.} \qquad \frac{\partial}{\partial x^\alpha} \left( \Box A^\alpha - \frac{\partial^2 A^\beta}{\partial x_\alpha \partial x^\beta} \right) = 0. \tag{9.55}$$

Equation (9.55) 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,

$$A'_\alpha = A_\alpha + \frac{\partial \Phi}{\partial x^\alpha}, \tag{9.56}$$

will also be a solution. In fact, by direct substitution we find

$$\Box A'_\alpha - \frac{\partial}{\partial x^\alpha} \Box \Phi - \eta^{\beta\delta} \frac{\partial^2 A'_\delta}{\partial x_\alpha \partial x^\beta} + \eta^{\beta\delta} \frac{\partial^2}{\partial x_\alpha \partial x^\beta} \frac{\partial \Phi}{\partial x^\delta} = -\frac{4\pi}{c} J_\alpha, \tag{9.57}$$

and since the second and the last term on the left hand-side cancel, it becomes

$$\Box A'_\alpha - \eta^{\beta\delta} \frac{\partial^2 A'_\delta}{\partial x_\alpha \partial x^\beta} = -\frac{4\pi}{c} J_\alpha, \tag{9.58}$$

q.e.d.

Since $\Phi$ is arbitrary, we can chose it in such a way that

$$\frac{\partial}{\partial x_\beta} A^{\beta\prime} = 0 \tag{9.59}$$

and eq. (9.58) becomes

$$\Box A'_\alpha - = -\frac{4\pi}{c} J_\alpha, \tag{9.60}$$

This is the Lorentz gauge.

Summaryzing: 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. (9.56). 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.

****************************************************************

## 9.4 Example: The armonic gauge.

****************************************************************

**The armonic gauge** is defined by the condition

$$\Gamma^{\lambda} = g^{\mu\nu}\Gamma^{\lambda}_{\mu\nu} = 0. \tag{9.61}$$

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, in fact given a generic coordinate transformation, the affine connections $\Gamma^{\alpha}_{\beta\gamma}$ transform as (see eq. (6.10))

$$\Gamma^{\lambda\prime}_{\mu\prime\nu\prime} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\frac{\partial x^{\tau}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}\Gamma^{\rho}_{\tau\sigma} - \frac{\partial x^{\rho}}{\partial x'^{\nu}}\frac{\partial x^{\sigma}}{\partial x'^{\mu}}\frac{\partial^2 x'^{\lambda}}{\partial x^{\rho}\partial x^{\sigma}}. \tag{9.62}$$

When contracted with $g'^{\mu\nu}$ they give

$$\Gamma^{\lambda\prime} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}}\Gamma^{\rho} - g^{\rho\sigma}\frac{\partial^2 x'^{\lambda}}{\partial x^{\rho}\partial x^{\sigma}}, \tag{9.63}$$

where we have made use of the equation

$$g^{\tau\sigma} = g^{\mu\prime\nu\prime}\frac{\partial x^{\tau}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}. \tag{9.64}$$

Therefore, if $\Gamma^\lambda$ is non zero, we can always find a frame where $\Gamma^{\rho\prime} = 0$ 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

$$\Gamma^\lambda = \frac{1}{2} g^{\mu\nu} g^{\lambda k} \left\{ \frac{\partial g_{k\mu}}{\partial x^\nu} + \frac{\partial g_{k\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^k} \right\} = 0. \tag{9.65}$$

Since

$$\begin{aligned} g^{\lambda k} \frac{\partial g_{k\mu}}{\partial x^\nu} &= -g_{k\mu} \frac{\partial g^{\lambda k}}{\partial x^\nu}, \\ \frac{1}{2} g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^k} &= \sqrt{-g} \frac{\partial}{\partial x^k} \frac{1}{\sqrt{-g}}, \end{aligned} \tag{9.66}$$

it follows that

$$\Gamma^\lambda = \frac{1}{2} g^{\mu\nu} \left\{ -g_{k\mu} \left[ \frac{\partial g^{\lambda k}}{\partial x^\nu} \right] - g_{k\nu} \left[ \frac{\partial g^{\lambda k}}{\partial x^\mu} \right] \right\} - \frac{g^{\lambda k}}{\sqrt{-g}} \frac{\partial}{\partial x^k} \sqrt{-g} = 0. \tag{9.67}$$

The term in brackets is symmetric in $\mu$ and $\nu$, therefore

$$\Gamma^\lambda = -\frac{1}{2} \left\{ 2 g^{\mu\sigma} g_{k\mu} \frac{\partial g^{\lambda k}}{\partial x^\sigma} \right\} - \frac{g^{\lambda k}}{\sqrt{-g}} \frac{\partial}{\partial x^k} \sqrt{-g} = 0, \tag{9.68}$$

and, since $g^{\mu\sigma} g_{k\mu} = \delta^\sigma{}_k$

$$\Gamma^\lambda = -\frac{\partial g^{\lambda k}}{\partial x^k} - \frac{g^{\lambda k}}{\sqrt{-g}} \frac{\partial}{\partial x^k} \sqrt{-g} = 0, \tag{9.69}$$

from which we find

$$-\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^k} \left( \sqrt{-g} g^{\lambda k} \right) = 0. \tag{9.70}$$

This means that

$$\Gamma^\lambda = 0 \qquad \text{implies} \qquad \frac{\partial}{\partial x^k} \left( \sqrt{-g} g^{\lambda k} \right) = 0. \tag{9.71}$$

The reason why this gauge is called 'armonic' is the following. A function $\Phi$ is armonic if

$$\Box\Phi = 0, \tag{9.72}$$

where the operator $\Box$ is the **covariant d'Alambertian operator** defined as

$$\Box\Phi = g^{\lambda k}\nabla_\lambda\nabla_k\Phi, \tag{9.73}$$

and $\nabla_\lambda$ is the covariant derivative. Since

$$g^{\lambda k}\nabla_\lambda\nabla_k\Phi = g^{\lambda k}\left(\frac{\partial\Phi_{;\lambda}}{\partial x^k} - \Gamma^\alpha_{\lambda k}\Phi_{;\alpha}\right) = \tag{9.74}$$
$$g^{\lambda k}\left[\frac{\partial^2\Phi}{\partial x^k\partial x^\lambda} - \Gamma^\alpha_{\lambda k}\frac{\partial\Phi}{\partial x^\alpha}\right] = g^{\lambda k}\frac{\partial^2\Phi}{\partial x^k\partial x^\lambda} - \Gamma^\alpha\frac{\partial\Phi}{\partial x^\alpha}.$$

If $\Gamma^\lambda = 0$ the armonic gauge condition becomes

$$\Box\Phi = g^{\lambda k}\frac{\partial^2\Phi}{\partial x^k\partial x^\lambda} = 0. \tag{9.75}$$

If $\Gamma^\lambda = 0$ then the **coordinates itself are armonic functions**, in fact putting $\Phi = x^\mu$ in eq. (9.75) one finds

$$\Box x^\mu = g^{\lambda k}\frac{\partial^2 x^\mu}{\partial x^k\partial x^\lambda} = g^{\lambda k}\frac{\partial}{\partial x^k}\delta^\mu_\lambda = 0, \tag{9.76}$$

q.e.d. If the spacetime is flat, armonic coordinates coincide with minkowskian coordinates.

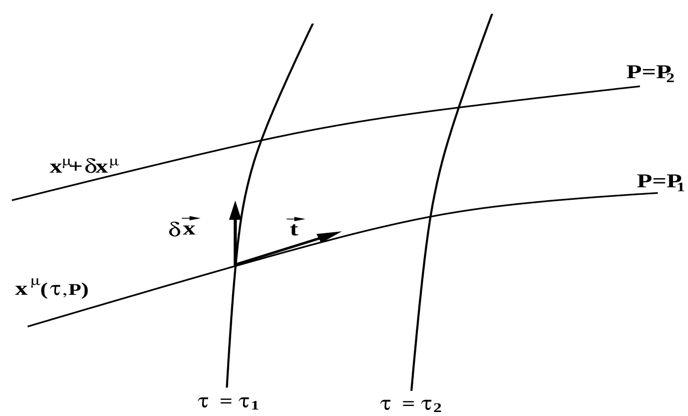
# Chapter 10

# 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. Conversely, 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.

## 10.1 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




Be

$$t^\alpha = \frac{\partial x^\alpha}{\partial \tau} \tag{10.1}$$

the tangent vector to the geodesic line, and be

$$\delta x^\alpha = \frac{\partial x^\alpha}{\partial p}. \tag{10.2}$$

Note that

$$\frac{\partial t^\alpha}{\partial p} = \frac{\partial \delta x^\alpha}{\partial \tau}. \tag{10.3}$$

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. $\nabla_{\vec{\delta x}}\, \vec{t} = \frac{d\vec{t}}{dp} = \frac{\partial \vec{t}}{\partial x^{\mu}}\frac{dx^{\mu}}{dp}$. The components of this vector are

$$\left(\nabla_{\vec{\delta x}}\, \vec{t}\right)^{\alpha} = \frac{\partial x^{\mu}}{\partial p}\left[\frac{\partial t^{\alpha}}{\partial x^{\mu}} + \Gamma^{\alpha}{}_{\mu\nu} t^{\nu}\right] = \frac{\partial t^{\alpha}}{\partial p} + \Gamma^{\alpha}{}_{\mu\nu} t^{\nu}\delta x^{\mu}. \tag{10.4}$$

The covariant derivative of the vector $\delta x$ along the curve $p = \text{const}$ i.e. along the geodesic, similarly has components

$$\left(\nabla_{\vec{t}}\, \vec{\delta x}\right)^{\alpha} = \left(\frac{d\vec{\delta x}}{d\tau}\right)^{\alpha} = t^{\mu}\delta x^{\alpha}{}_{;\mu} = \frac{\partial \delta x^{\alpha}}{\partial \tau} + \Gamma^{\alpha}{}_{\mu\nu}\delta x^{\nu} t^{\mu}. \tag{10.5}$$

From eq. (10.3) and from the symmetry of $\Gamma^{\alpha}{}_{\mu\nu}$ in the lower indices it follows that

$$\nabla_{\vec{t}}\, \vec{\delta x} = \nabla_{\vec{\delta x}}\, \vec{t}. \tag{10.6}$$

The quantities $\left(\nabla_{\vec{t}}\, \vec{\delta x}\right)^{\alpha}$ or $\left(\nabla_{\vec{\delta x}}\, \vec{t}\right)^{\alpha}$ 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 $\vec{\delta x}$ along the curve $p = const$, i.e $\nabla_{\vec{t}}\left(\nabla_{\vec{t}}\, \vec{\delta x}\right)$. From eq. (10.6) it follows that

$$\nabla_{\vec{t}}\left(\nabla_{\vec{t}}\, \vec{\delta x}\right) = \nabla_{\vec{t}}\left(\nabla_{\vec{\delta x}}\, \vec{t}\right). \tag{10.7}$$

This equation can be rewritten as

$$\frac{d^2 \delta x^{\alpha}}{d\tau^2} = \left[\nabla_{\vec{t}}\left(\nabla_{\vec{\delta x}}\, \vec{t}\right)\right]^{\alpha}. \tag{10.8}$$

In order to develop the right-hand side of eq. (10.8), consider $\nabla_{\vec{\delta x}} \vec{t}$ as a vector $\vec{A}$

$$A^\alpha = \left(\nabla_{\vec{\delta x}} \vec{t}\right)^\alpha = \frac{\partial t^\alpha}{\partial p} + \Gamma^\alpha{}_{\mu\nu} t^\mu \delta x^\nu, \tag{10.9}$$

and apply the operator $\nabla_{\vec{t}}$ :

$$\left(\nabla_{\vec{t}} \vec{A}\right)^\alpha = \frac{\partial A^\alpha}{\partial \tau} + \Gamma^\alpha{}_{\rho\nu} A^\rho t^\nu. \tag{10.10}$$

Following this procedure we find

$$\begin{aligned} \frac{d^2 \delta x^\alpha}{d\tau^2} &= \frac{\partial^2 t^\alpha}{\partial \tau \partial p} + \Gamma^\alpha{}_{\mu\nu,\beta} t^\beta t^\mu \delta x^\nu \\ &+ \Gamma^\alpha{}_{\mu\nu} \left( \frac{\partial t^\mu}{\partial \tau} \delta x^\nu + t^\mu \frac{\partial \delta x^\nu}{\partial \tau} \right) + \Gamma^\alpha{}_{\rho\gamma} \left( \frac{\partial t^\rho}{\partial p} + \Gamma^\rho{}_{\mu\nu} t^\mu \delta x^\nu \right) t^\gamma. \end{aligned} \tag{10.11}$$

This equation can be further simplified. In fact, since $t^\mu$ is the geodesic tangent vector, when it is parallel-transported along the geodesic it gives (see chapter 6.2)

$$\nabla_{\vec{t}} \vec{t} = 0, \tag{10.12}$$

and consequently

$$\nabla_{\vec{\delta x}} \left(\nabla_{\vec{t}} \vec{t}\right) = 0. \tag{10.13}$$

This derivative can be computed by using the same procedure described above: introduce a vector $\vec{B}$ whose components are $B^\alpha = \left(\nabla_{\vec{t}} \vec{t}\right)^\alpha = \frac{\partial t^\alpha}{\partial \tau} + \Gamma^\alpha{}_{\mu\nu} t^\mu t^\nu$, and take its directional derivative along the geodesic. Eq. (10.13) then becomes

$$\frac{\partial^2 t^\alpha}{\partial \tau \partial p} = -\Gamma^\alpha{}_{\mu\nu,\beta} t^\mu t^\nu \delta x^\beta \tag{10.14}$$

$$-\Gamma^{\alpha}{}_{\mu\nu}\left(\frac{\partial t^{\mu}}{\partial p}t^{\nu}+\frac{\partial t^{\nu}}{\partial p}t^{\mu}\right)-\Gamma^{\alpha}{}_{\rho\gamma}\left(\frac{\partial t^{\rho}}{\partial \tau}+\Gamma^{\rho}{}_{\mu\nu}t^{\mu}t^{\nu}\right)\delta x^{\gamma}.$$

If we now substitute eq. (10.14) in eq. (10.11), and use the symmetries of the $\Gamma^{\alpha}{}_{\mu\nu}$ , we find

$$\frac{d^2\delta x^{\alpha}}{d\tau^2}=t^{\beta}t^{\mu}\delta x^{\nu}\left(\Gamma^{\alpha}{}_{\mu\nu,\beta}-\Gamma^{\alpha}{}_{\mu\beta,\nu}+\Gamma^{\alpha}{}_{\rho\beta}\Gamma^{\rho}{}_{\mu\nu}-\Gamma^{\alpha}{}_{\rho\nu}\Gamma^{\rho}{}_{\mu\beta}\right). \tag{10.15}$$

The term in parenthesis is the Riemann tensor, therefore

$$\frac{d^2\delta x^{\alpha}}{d\tau^2}=R^{\alpha}{}_{\mu\beta\nu}\frac{dx^{\beta}}{d\tau}\frac{dx^{\mu}}{d\tau}\delta x^{\nu}. \tag{10.16}$$

This is the equation of the **geodesic deviation**. Since the Riemann tensor appears on the right-hand side of this equation, and since it is zero if and only if the gravitational field either zero or constant and uniform, the equation of the geodesic deviation really contains meaningfull information on the gravitational field.

# Chapter 11

# 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 we can think that when a distribution of masses evolves in the spacetime, the information about this motion should propagate in the form of waves. Since both the gravitational potential and the spacetime metric are expressed by the metric tensor $g_{\mu\nu}$, gravitational waves are **metric waves**. When they propagate through the space, the geometry, and consequently the distance between points, changes in time.

Gravitational waves can be studied by following two different approaches

1) Assume that a given metric, which is an exact solution of Einstein's equations, is perturbed by some external agency. This perturbation induces small changes in the gravitational field, and the corresponding metric tensor can be written as

$$g_{\mu\nu} = g^0_{\mu\nu} + h_{\mu\nu}, \tag{11.1}$$

where the superscript zero indicates the unperturbed exact solution, (for example the flat metric $\eta_{\mu\nu}$ or the metric generated by a black hole or a star), and $h_{\mu\nu}$ is a small perturbation

$$|h_{\mu\nu}| << 1.$$

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 reference does exists.

The Einstein tensor computed for the metric (11.1) contains non-linear terms involving products of first derivatives of $g_{\mu\nu}$. These terms are second order in $h_{\mu\nu}$, and if we restrict our study to first order perturbations they can be neglected. Then the resulting equations will be linear, and their solution will describe the propagation of gravitational waves in the considered background. This approximation works sufficiently well in a large variety of physical situations, because gravitational waves are very weak. This point will be better understood when we shall study the generation of gravitational waves.

2) A different approach is to look for **exact solutions** of Einstein's equations. The problem of finding an exact solution which describes both the source and the emitted wave is still unsolved. Of course the non-linearity of the equations makes the problem very difficult, but it should be remembered that in electromagnetism also an exact solution describing the electromagnetic field produced by a current that decreases due to the emission of waves in an electric oscillator has never been found.

Exact solutions of Einstein's equations which describe gravitational waves propagating in vacuum 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.

## 11.1 Perturbations of a flat spacetime

In this section we shall start the study of gravitational waves by considering the simplest case: very weak disturbances propagating in flat spacetime. We will use the perturbative approach, and solve the linearized Einstein equations.

Consider a flat spacetime $\eta_{\mu\nu}$ and a small perturbation propagating through it, such that the resulting metric can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| << 1. \tag{11.2}$$

Since we are interested only in first order terms, we shall raise and lower indices with the flat metric $\eta^{\mu\nu}$. Let us write Einstein's equations in the form

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{11.3}$$

Remember that

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \qquad R_{\mu k} = g^{\lambda\nu} R_{\lambda\mu\nu k} = R^{\nu}{}_{\mu\nu k}, \tag{11.4}$$

and that the Riemann tensor is

$$R^{\lambda}{}_{\mu\nu k} = \frac{\partial \Gamma^{\lambda}{}_{\mu k}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\lambda}{}_{\mu\nu}}{\partial x^{k}} + \Gamma^{\eta}{}_{\mu k}\Gamma^{\lambda}{}_{\nu\eta} - \Gamma^{\eta}{}_{\mu\nu}\Gamma^{\lambda}{}_{k\eta}. \tag{11.5}$$

Consequently the Ricci tensor is

$$R_{\mu k} = R^{\nu}{}_{\mu\nu k} = \frac{\partial \Gamma^{\nu}{}_{\mu k}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\nu}{}_{\mu\nu}}{\partial x^{k}} + \Gamma^{\eta}{}_{\mu k}\Gamma^{\nu}{}_{\nu\eta} - \Gamma^{\eta}{}_{\mu\nu}\Gamma^{\nu}{}_{k\eta}. \tag{11.6}$$

By retaining only first order terms in $h_{\mu\nu}$, it becomes

$$R_{\mu k} \sim \frac{\partial \Gamma^{\nu}{}_{\mu k}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\nu}{}_{\mu\nu}}{\partial x^{k}} + O(h^2). \tag{11.7}$$

The affine connections computed for the metric (11.2) are

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2}\eta^{\lambda\rho}\left[\frac{\partial}{\partial x^{\mu}} h_{\rho\nu} + \frac{\partial}{\partial x^{\nu}} h_{\rho\mu} - \frac{\partial}{\partial x^{\rho}} h_{\mu\nu}\right] + O(h^2). \tag{11.8}$$

By using these equations we find that

$$G_{\mu\nu} \sim \frac{1}{2}\left\{-\Box_F h_{\mu\nu} + \left[\frac{\partial^2}{\partial x^\lambda \partial x^\mu} h^\lambda_\nu + \frac{\partial^2}{\partial x^\lambda \partial x^\nu} h^\lambda_\mu - \frac{\partial^2}{\partial x^\mu \partial x^\nu} h^\lambda_\lambda\right]\right\}, \qquad (11.9)$$

where $\Box_F = -\frac{\partial^2}{c^2 \partial t^2} + \nabla^2$. [1]

In eq. (11.3) $T_{\mu\nu}$ is the stress-energy tensor associated to the source of the perturbation and it satisfies the conservation equation

$$T^{\mu\nu}{}_{;\nu} = 0,$$

which, in the weak field limit, becomes

$$\frac{\partial}{\partial x^\mu} T^{\mu\nu} = 0.$$

[1] If one wants to compare our expression of $G_{\mu\nu}$ with eq. 8.32 given in Schutz:

$$G_{\mu\nu} = -\frac{1}{2}\left[h_{\mu\nu,\alpha}{}^{,\alpha} + \eta_{\mu\nu} h_{\alpha\beta}{}^{,\alpha,\beta} - h_{\mu\alpha,\nu}{}^{,\alpha} - h_{\nu\alpha,\mu}{}^{,\alpha}\right],$$

it is immediate to check that the first term is the d'Alambertian

$$h_{\mu\nu,\alpha}{}^{,\alpha} = \eta^{\alpha\beta} h_{\mu\nu,\alpha,\beta}.$$

Consider now the terms in square bracket in eq. (11.9). They can be rewritten in the following form

$$\begin{aligned}
&\left[\frac{\partial^2}{\partial x^\lambda \partial x^\mu} h^\lambda_\nu + \frac{\partial^2}{\partial x^\lambda \partial x^\nu} h^\lambda_\mu - \frac{\partial^2}{\partial x^\mu \partial x^\nu} h^\lambda_\lambda\right] = \\
&= \eta^{\lambda\alpha} h_{\alpha\nu,\lambda\mu} + \eta^{\lambda\alpha} h_{\alpha\mu,\lambda\nu} - \eta^{\lambda\alpha} h_{\alpha\lambda,\mu\nu} = \\
&= h_{\nu\alpha,\mu}{}^{,\alpha} + h_{\mu\alpha,\nu}{}^{,\alpha} - \eta_{\nu\alpha}\eta_{\lambda\mu}\eta^{\lambda\alpha} h_{\alpha\lambda}{}^{,\alpha,\lambda}
\end{aligned}$$

that coincide with Schutz's terms.

Einstein's equations finally are

$$\left\{\Box_F h_{\mu\nu} - \left[\frac{\partial^2}{\partial x^\lambda \partial x^\mu} h^\lambda_\nu + \frac{\partial^2}{\partial x^\lambda \partial x^\nu} h^\lambda_\mu - \frac{\partial^2}{\partial x^\mu \partial x^\nu} h^\lambda_\lambda\right]\right\} = -\frac{16\pi G}{c^4} T_{\mu\nu}. \quad (11.10)$$

In order to take into account the back reaction due to the energy and momentum transported by the gravitational waves traveling in the flat background, one should add on the right-hand-side the energy-momentum pseudotensor of the gravitational field $t_{\mu\nu}$. $t_{\mu\nu}$ can be constructed in a variety of ways by imposing the condition that, when added to $T_{\mu\nu}$ it satisfies the conservation law

$$\left(t^{\mu\nu} + T^{\mu\nu}\right)_{;\nu} = 0.$$

The definition of $t_{\mu\nu}$ would require a detailed discussion on how to construct conservation laws for perturbed spacetimes, which is a very interesting and conceptually relevant problem. However in what follows we shall restrict our study only to the solution of eq. (11.10).

As already discussed in chapter 9, the solution of eqs. (11.10) is not uniquely determined. If we make a coordinate transformation, the transformed metric tensor will still be a solution: it will describe 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 $|h'_{\mu\nu}| << 1$ (here and in the following we shall use the notation $h'_{\mu\nu} \equiv h_{\mu'\nu'}$). If

$$x^{\mu\prime} = x^\mu + \epsilon^\mu(x), \quad (11.11)$$

where $\epsilon^\mu$ is an arbitrary vector such that $\frac{\partial \epsilon^\mu}{\partial x^\nu}$ is of the same order of $h_{\mu\nu}$, it is easy to check that since

$$g'_{\mu\nu} = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial x^{\mu\prime}} \frac{\partial x^\beta}{\partial x^{\nu\prime}} = (\eta_{\alpha\beta} + h_{\alpha\beta}) \frac{\partial x^\alpha}{\partial x^{\mu\prime}} \frac{\partial x^\beta}{\partial x^{\nu\prime}}, \tag{11.12}$$

then

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \epsilon_\mu}{\partial x^\nu} - \frac{\partial \epsilon_\nu}{\partial x^\mu}. \tag{11.13}$$

In order to simplify eq. (11.10) it appears convenient to choose a coordinate system in which the harmonic gauge condition is satisfied

$$g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = 0. \tag{11.14}$$

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 [2]

$$\frac{\partial}{\partial x^\mu} h^\mu{}_\nu = \frac{1}{2} \frac{\partial}{\partial x^\nu} h^\mu{}_\mu. \tag{11.15}$$

If this is true, it is clear that the term in square brackets in eq. (11.10) vanishes, and it reduces to a simple wave equation supplemented by the

[2]

$$g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = \frac{1}{2} \eta^{\mu\nu} \eta^{\lambda k} \left\{ \frac{\partial h_{k\mu}}{\partial x^\nu} + \frac{\partial h_{k\nu}}{\partial x^\mu} - \frac{\partial h_{\mu\nu}}{\partial x^k} \right\} = \frac{1}{2} \eta^{\lambda k} \left\{ h^\nu{}_{k,\nu} + h^\mu{}_{k,\mu} - h^\nu{}_{\nu,k} \right\}$$

Since the first two terms are equal we find

$$0 = g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = \eta^{\lambda k} \left\{ h^\mu{}_{k,\mu} - \frac{1}{2} h^\nu{}_{\nu,k} \right\}$$

q.e.d.

condition (11.15)

$$\begin{cases} \Box_F h_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}. \\ \frac{\partial}{\partial x^\mu} h^\mu{}_\nu = \frac{1}{2} \frac{\partial}{\partial x^\nu} h^\mu{}_\mu. \end{cases} \tag{11.16}$$

In vacuum they become

$$\begin{cases} \Box_F h_{\mu\nu} = 0. \\ \frac{\partial}{\partial x^\mu} h^\mu{}_\nu = \frac{1}{2} \frac{\partial}{\partial x^\nu} h^\mu{}_\mu. \end{cases} \tag{11.17}$$

**Thus, we have shown that a perturbation of a flat spacetime propagates as a wave at the speed of light, and that Einstein's theory of gravity predicts the existence of gravitational waves**.

In appendix 11A we shall show how to choose $\epsilon^\mu$ in such a way that the condition of harmonic gauge is satisfied.

The solution of eqs. (11.16) can be written in terms of retarded potential as in electrodynamics

$$h_{\mu\nu}(\mathbf{x}, t) = \frac{4G}{c^4} \int \frac{d^3x' T_{\mu\nu}(\mathbf{x}', t - |\mathbf{x}\text{-}\mathbf{x}'|)}{|\mathbf{x}\text{-}\mathbf{x}'|}, \tag{11.18}$$

and the integral extends over the past light-cone of the event $(t, \mathbf{x})$. This equation represents the gravitational waves generated by the source $T_{\mu\nu}$.

We may now ask how eqs. (11.17) and (11.16) should be modified if, in place of considering a disturbance propagating in flat spacetime, we would consider a perturbation of a curved background. For example, suppose we know the exact solution $g^0_{\mu\nu}$ which describes the gravitational field external to a spherically symmetric object, a star or a black hole, If $h_{\mu\nu}$ is a perturbation of $g^0_{\mu\nu}$ as described in eq. (11.1), it is possible to show that, by a suitable choice of the gauge, the Einstein equations written for certain combinations of the metric function, say $\Phi$, can be reduced to a form similar to eqs. (11.16) or (11.17). 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 the waves are scattered and the final equation will have the form

$$\Box_F \Phi - V(x^\mu)\Phi = -\frac{16\pi G}{c^4} T_{\mu\nu}, \tag{11.19}$$

where $\Box_F$ 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 sperically symmetric, stationary gravitational field would be described by a Schroedinger-like equation! A complete account on the theory of perturbations of black holes is given in the book *The Mathematical Theory of Black Holes* by S. Chandrasekhar, Oxford: Claredon Press, (1984).

## 11.2 Plane gravitational waves

The simplest solution of the wave equation

$$\Box_F h_{\mu\nu} = 0, \tag{11.20}$$

is a monocromatic plane wave

$$h_{\mu\nu} = \Re\left\{ A_{\mu\nu} e^{ik_\alpha x^\alpha} \right\}, \tag{11.21}$$

where $A_{\mu\nu}$ is the amplitude of the wave, called the polarization tensor, and $\vec{k}$ is the wave-vector. Let us verify if eq.(11.21) is really a solution. By direct substitution into eq. (11.20) we find

$$\Box_F h_{\mu\nu} = \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\left(e^{ik_\gamma x^\gamma}\right) = \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\left[ik_\gamma\frac{\partial x^\gamma}{\partial x^\beta}e^{ik_\gamma x^\gamma}\right] = \tag{11.22}$$

$$\eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\left[ik_\gamma\delta^\gamma{}_\beta e^{ik_\gamma x^\gamma}\right] = \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\left[ik_\beta e^{ik_\gamma x^\gamma}\right] =$$

$$= -\eta^{\alpha\beta}k_\alpha k_\beta = 0.$$

It follows that $\vec{k}$ is a null vector. In addition the harmonic gauge condition requires that

$$\frac{\partial}{\partial x^\mu}h^\mu{}_\nu = \frac{1}{2}\frac{\partial}{\partial x^\nu}h^\mu{}_\mu, \tag{11.23}$$

which can be written as

$$\eta^{\mu\alpha}\frac{\partial}{\partial x^\mu}h_{\alpha\nu} = \frac{1}{2}\eta^{\mu\beta}\frac{\partial}{\partial x^\nu}h_{\beta\mu}. \tag{11.24}$$

Using eq. (11.21) it becomes

$$\eta^{\mu\alpha}\frac{\partial}{\partial x^\mu}A_{\alpha\nu}e^{ik_\gamma x^\gamma} = \frac{1}{2}\eta^{\mu\beta}\frac{\partial}{\partial x^\nu}A_{\beta\mu}e^{ik_\delta x^\delta}, \tag{11.25}$$

i.e.

$$\eta^{\mu\alpha}A_{\alpha\nu}k_\mu = \frac{1}{2}\eta^{\mu\beta}A_{\beta\mu}k_\nu \tag{11.26}$$

and finally

$$k_\mu A^\mu{}_\nu = \frac{1}{2}k_\nu A^\mu{}_\mu. \tag{11.27}$$

Thus eq. (11.21) is a solution of eq. (11.20) if $\vec{k}$ is a **null vector**, and if the condition (11.27) is satisfied.

Since $h_{\mu\nu}$ is constant on those surfaces where

$$k_\alpha x^\alpha = 0, \tag{11.28}$$

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

$$\vec{k} \to (\omega, \mathbf{k}). \tag{11.29}$$

Since $\vec{k}$ is a null vector

$$-(k_0)^2 + (k_x)^2 + (k_y)^2 + (k_z)^2 = 0, \qquad \text{i.e.} \tag{11.30}$$

$$\omega = ck_0 = c\sqrt{(k_x)^2 + (k_y)^2 + (k_z)^2}. \tag{11.31}$$

This is the general case of a wave which propagates in the direction

$$\frac{1}{k^0}(k_x, k_y, k_z), \tag{11.32}$$

## 11.3 The TT-gauge

We now want to see which one of the ten components of $h_{\mu\nu}$ (or of $A_{\mu\nu}$ in the case of plane waves) has a real physical meaning. Let us consider a plane gravitational wave propagating in flat spacetime along the $x^1 = x$-direction. Since $h_{\mu\nu}$ is independent on y and z, eqs. (11.17) become (as usual we raise and lower indices with $\eta_{\mu\nu}$ )

$$\left(-\frac{\partial^2}{c^2\partial t^2}+\frac{\partial^2}{\partial x^2}\right)h^{\mu}{}_{\nu}=0, \tag{11.33}$$

$$\frac{\partial}{\partial x^{\mu}}h^{\mu}{}_{\nu}=\frac{1}{2}\frac{\partial}{\partial x^{\nu}}h^{\mu}{}_{\mu}. \tag{11.34}$$

Consequently, $h^{\mu}{}_{\nu}$ is an arbitrary function of $t\pm\frac{x}{c}$ . Equation (11.34) can be rewritten in the alternative form

$$\frac{\partial}{\partial x^{\mu}}\left[h^{\mu}{}_{\nu}-\frac{1}{2}\delta^{\mu}{}_{\nu}h^{\alpha}{}_{\alpha}\right]=0, \tag{11.35}$$

and if we introduce the tensor

$$\Psi^{\mu\nu}=\left[h^{\mu\nu}-\frac{1}{2}\eta^{\mu\nu}h^{\alpha}{}_{\alpha}\right]. \tag{11.36}$$

eq. (11.35) becomes

$$\frac{\partial}{\partial x^{\mu}}\Psi^{\mu}{}_{\nu}=0. \tag{11.37}$$

This equation, equivalent to eq. (11.34), expresses the harmonic gauge condition (see eq. (11.14)). If we restrict our analysis to progressive waves, $\Psi^{\mu}{}_{\nu}$ will depend only on $t-\frac{x}{c}=\chi(t,x)$, i.e.

$$\Psi^{\mu}{}_{\nu}=\Psi^{\mu}{}_{\nu}\left[\chi(t,x)\right] \tag{11.38}$$

and since

$$\begin{cases} \frac{\partial}{\partial t}\Psi^{\mu}{}_{\nu}=\frac{\partial\Psi^{\mu}{}_{\nu}}{\partial\chi}\frac{\partial\chi}{\partial t}=\frac{\partial\Psi^{\mu}{}_{\nu}}{\partial\chi}, \\ \frac{\partial}{\partial x}\Psi^{\mu}{}_{\nu}=\frac{\partial\Psi^{\mu}{}_{\nu}}{\partial\chi}\frac{\partial\chi}{\partial x}=-\frac{1}{c}\frac{\partial\Psi^{\mu}{}_{\nu}}{\partial\chi}, \end{cases} \tag{11.39}$$

eq. (11.37) gives

$$\frac{\partial}{\partial x^{\mu}}\Psi^{\mu}{}_{\nu}=\frac{1}{c}\frac{\partial\Psi^{t}{}_{\nu}}{\partial t}+\frac{\partial\Psi^{x}{}_{\nu}}{\partial x}=\frac{1}{c}\frac{\partial}{\partial\chi}\left[\Psi^{t}{}_{\nu}-\Psi^{x}{}_{\nu}\right]=0. \tag{11.40}$$

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 of the integration is

$$\Psi^t{}_t = \Psi^x{}_t, \qquad \Psi^t{}_y = \Psi^x{}_y, \tag{11.41}$$
$$\Psi^t{}_x = \Psi^x{}_x, \qquad \Psi^t{}_z = \Psi^x{}_z.$$

We now observe that the harmonic gauge condition (11.37) does not determine the gauge uniquely. In fact we still have the freedom of making an infinitesimal coordinate transformation

$$x'^\mu = x^\mu + \xi^\mu(t - \frac{x}{c}), \tag{11.42}$$

provided $\xi^\mu$ satisfies the wave equation

$$\Box_F \xi^\mu(t - \frac{x}{c}) = 0, \tag{11.43}$$

and we can use the four functions $\xi^\mu(t - \frac{x}{c})$ to set to zero the following four quantities

$$\Psi^t{}_x = \Psi^t{}_y = \Psi^t{}_z = \Psi^y{}_y + \Psi^z{}_z = 0. \tag{11.44}$$

From eq. (11.41) it then follows that

$$\Psi^x{}_x = \Psi^x{}_y = \Psi^x{}_z = \Psi^t{}_t = 0, \tag{11.45}$$

where $\Psi^t{}_t$ is zero because

$$\Psi^x{}_t = -\Psi^t{}_x. \tag{11.46}$$

The remaining non-vanishing components are $\Psi^z{}_y$ and $\Psi^y{}_y - \Psi^z{}_z$. These components cannot be set equal to zero, because we have exhausted our gauge freedom.

From eqs. (11.44) and (11.45) it follows that

$$\Psi^\mu{}_\mu = \Psi^t{}_t + \Psi^x{}_x + \Psi^y{}_y + \Psi^z{}_z = 0, \tag{11.47}$$

and since

$$\Psi^\mu{}_\mu = h^\mu{}_\mu - \frac{1}{2} h^\mu{}_\mu, \tag{11.48}$$

it follows that

$$h^\mu{}_\mu = 0, a \qquad \rightarrow \qquad \Psi^\mu{}_\nu = h^\mu{}_\nu, \tag{11.49}$$

both $h_{\mu\nu}$ and $\Psi_{\mu\nu}$ are traceless. In conclusion a plane gravitational wave propagating along the $x$-axis is characterized by two functions $h_{23} = h_{xy}$ e $h_{22} = -h_{33}$ $(h_{yy} = -h_{zz})$, and the remaining components can be set to zero by a suitable choice of the gauge:

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & h_{yy} & h_{yz} \\ 0 & 0 & h_{yz} & -h_{yy} \end{pmatrix}. \tag{11.50}$$

**Thus, a gravitational wave has only two physical degrees of freedom which correspond to the two possible states of polarization.** The gauge in which this is clearly manifested is called the TT-gauge, where 'TT-'

indicates that the metric perturbation $h_{\mu\nu}$ is different from zero only on the plane orthogonal to the direction of propagation (transverse), and that it is traceless.

## 11.4 How to find the TT-gauge

Consider a plane wave solution of the wave equation eq. (11.20)

$$h_{\mu\nu} = A_{\mu\nu} e^{ik_\alpha x^\alpha}. \tag{11.51}$$

Here we have omitted $\Re$, but it should be understood that at the end we will take the real part of the quantities we shall compute. Choose $\xi_\alpha$ of this form

$$\xi_\alpha = B_\alpha e^{ik_\alpha x^\alpha} \tag{11.52}$$

where $k^\mu$ is the same null vector as for the solution (11.51). We know $A_{\mu\nu}$, and we want to determine $B_\alpha$ in such a way that the harmonic gauge condition is satisfied.

In the new frame

$$h'_{\alpha\beta} = h_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha}, \tag{11.53}$$

$$\Psi'_{\alpha\beta} = \left[h'_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta} h'^{\mu}{}_{\mu}\right] = \Psi_{\alpha\beta} - \xi_{\alpha,\beta} - \xi_{\beta,\alpha} + \eta_{\alpha\beta}\xi^{\mu}{}_{,\mu}. \tag{11.54}$$

We now substitute eqs. (11.51) and (11.52) in eq. (11.54), impose that in the new gauge $\Psi'_{\alpha\beta} = h'_{\alpha\beta}$ (see eq. (11.49), and find

$$A'_{\alpha\beta} = A_{\alpha\beta} - iB_\alpha k_\beta - iB_\beta k_\alpha + i\eta_{\alpha\beta} B^\mu k_\mu. \tag{11.55}$$

We now impose the TT-gauge conditions

1) *traceless condition*

In order to satisfy this condition we multiply eq. (11.55) by $\eta^{\alpha\beta}$ and find

$$A'^{\alpha}{}_{\alpha} = A^{\alpha}{}_{\alpha} - iB^{\beta}k_{\beta} - iB^{\alpha}k_{\alpha} + iB^{\mu}k_{\mu} = A^{\alpha}{}_{\alpha} - iB^{\beta}k_{\beta} = 0. \tag{11.56}$$

This equation provides one condition on $B_{\alpha}$.

2) *"Transverse"-condition*

Choose an arbitrary and constant timelike vector $U^{\alpha}$ (for example we may choose $U^{\beta} = \delta^{\beta}{}_{0}$), and impose the following condition

$$A'_{\alpha\beta}U^{\beta} = A'_{\alpha 0} = 0. \tag{11.57}$$

i.e.

$$A_{\alpha 0} - iB_{\alpha}k_0 - iB_0 k_{\alpha} - iB^{\mu}k_{\mu} = 0. \tag{11.58}$$

This equation can be used to put

$$A'_{x0} = A'_{yo} = A'_{z0} = 0. \tag{11.59}$$

If we now orient the axes of our frame in such a way that the incoming wave travels along the x-direction, from the harmonic gauge condition $\Psi^{\mu}_{\nu}$ satisfies, $\frac{\partial \Psi^{\mu}_{\nu}}{\partial x^{\mu}} = 0$, it follows that (compare with eq. (11.41)

$$A'_{00} = 0, \quad A'_{xx} = 0, \quad A'_{xy} = 0, \quad A'_{xz} = 0, \tag{11.60}$$

while eq. (11.56) implies that

$$A'_{yy} = -A'_{zz}. \tag{11.61}$$

Thus we have reduced the wave to the transverse-traceless form (remember that we have assumed that the wave travels along the x-axis).

## 11.5 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 the incoming wave. We can also set a $TT$-gauge as explained earlier (for example the four-velocity $U^\alpha$ may coincide with the initial four-velocity of the particle). When the wave arrives, the particle will follow the geodesic equation

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha{}_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \equiv \frac{dU^\alpha}{d\tau} + \Gamma^\alpha{}_{\mu\nu}U^\mu U^\nu = 0. \tag{11.62}$$

At $t = 0$ the particle is at rest and

$$\left(\frac{dU^\alpha}{d\tau}\right)_{(t=0)} = -\Gamma^\alpha{}_{00} = -\frac{1}{2}\eta^{\alpha\beta}\left[h_{\beta 0,0} + h_{0\beta,0} + h_{00,\beta}\right]. \tag{11.63}$$

Since we are in the $TT$-gauge it follows that

$$\left(\frac{dU^\alpha}{d\tau}\right)_{(t=0)} = 0. \tag{11.64}$$

Thus, the particle is not accelerated neither at $t = 0$ nor later! It remains at a **constant coordinate position**, regardeless of the wave, and the same will be true for any particle which is at rest before the passage of the wave. We conclude that **the study of the motion of a single particle does not allow to detect the passage of a gravitational wave.**

## 11.6 Geodesic deviation induced by a gravitational wave

We shall now study the relative motion of particles in the gravitational field produced by a gravitational wave.

Consider two neighbouring particles A and B, and a $TT$-reference frame constructed as explained in Section 4, and with the x-axis aligned with the direction of propagation of the incoming wave. If we choose the origin coincident with the position of the particle A

$$x_A^\lambda = (1, 0, 0, 0). \tag{11.65}$$

$\tau = ct$ is the proper time of the particle A. Since the two particles are initially at rest, they will remain at a constant coordinate position even later, when the wave arrives. However, since the metric changes, the proper distance between them will change. For example if the particle B is initially at some point on the y-axis

$$\Delta l = \int |ds^2|^{\frac{1}{2}} = \int_0^{y_B} |g_{yy}|^{\frac{1}{2}} dy = \int_0^{y_B} |1 + h^{TT}{}_{yy}|^{\frac{1}{2}} dy \neq \text{constant}. \quad (11.66)$$

Another way of studying the effect of the passage of the wave, is by means of the equation of the geodesic deviation. Be $\delta x^\mu$ the vector which separates the two particles, i.e. initially

$$\delta x^\mu = (0, x_B, y_B, z_B).$$

The equation of geodesic deviation found in chapter 10, written in the gauge we have choosen, is

$$\frac{d^2}{d\tau^2}\delta x^\lambda = R^\lambda{}_{00\mu}\delta x^\mu \quad (11.67)$$

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

$$\begin{aligned} R_{iklm} &= \frac{1}{2}\left(\frac{\partial^2 g_{im}}{\partial x^k \partial x^l} + \frac{\partial^2 g_{kl}}{\partial x^i \partial x^m} - \frac{\partial^2 g_{il}}{\partial x^k \partial x^m} - \frac{\partial^2 g_{km}}{\partial x^i \partial x^l}\right) + \quad (11.68) \\ &+ g_{np}\left(\Gamma^n{}_{kl}\Gamma^p{}_{im} - \Gamma^n{}_{km}\Gamma^p{}_{il}\right), \end{aligned}$$

after neglecting terms which are second order in $h_{\mu\nu}$, becomes

$$R_{iklm} = \frac{1}{2}\left(\frac{\partial^2 h_{im}}{\partial x^k \partial x^l} + \frac{\partial^2 h_{kl}}{\partial x^i \partial x^m} - \frac{\partial^2 h_{il}}{\partial x^k \partial x^m} - \frac{\partial^2 h_{km}}{\partial x^i \partial x^l}\right) + O(h^2), \quad (11.69)$$

and consequently

$$R_{i00m} = \frac{1}{2}\left(\frac{\partial^2 h_{im}}{\partial x^0 \partial x^0} + \frac{\partial^2 h_{00}}{\partial x^i \partial x^m} - \frac{\partial^2 h_{i0}}{\partial x^0 \partial x^m} - \frac{\partial^2 h_{0m}}{\partial x^i \partial x^0}\right). \quad (11.70)$$

Since we are in the $TT$-gauge, $h_{i0} = h_{00} = 0$ and it simply becomes

$$R_{i00m} = \frac{1}{2} h^{TT}_{im,00}, \tag{11.71}$$

where $i$ and $m$ can assume only the values 2 and 3, i.e. they refer to the $y$ and $z$ components. It follows that

$$R^{\lambda}{}_{00m} = \eta^{\lambda i} R_{i00m} = \frac{1}{2} \frac{\partial^2 h^{TT\lambda}{}_m}{c^2 \partial t^2}, \tag{11.72}$$

and the equation of geodesic deviation (11.67) becomes

$$\frac{d^2}{dt^2} \delta x^{\lambda} = \frac{1}{2} \frac{\partial^2 h^{TT\lambda}{}_m}{\partial t^2} \delta x^m. \tag{11.73}$$

In this equation the transverse nature of the wave is clearly manifested. In fact, since $h^{TT\lambda}{}_k$ is non-zero only in the $(y-z)$-plane, orthogonal to the direction of propagation, the particles will be accelerated on that plane, i.e.

$$\begin{aligned}
&\tfrac{d^2}{dt^2}\delta x^0 = \tfrac{d^2}{dt^2}\delta x^1 = 0 && \rightarrow \quad \delta x^0 = \delta x^1 = \text{const.} \\
&2\tfrac{d^2}{dt^2}\delta x^2 = \tfrac{\partial^2 h^{TT2}{}_2}{\partial t^2}\delta x^2 + \tfrac{\partial^2 h^{TT2}{}_3}{\partial t^2}\delta x^3 && \neq 0 \\
&2\tfrac{d^2}{dt^2}\delta x^3 = \tfrac{\partial^2 h^{TT3}{}_2}{\partial t^2}\delta x^2 + \tfrac{\partial^2 h^{TT3}{}_3}{\partial t^2}\delta x^3 && \neq 0.
\end{aligned} \tag{11.74}$$

Suppose that at $t = 0$ the two particles are at rest relative to each other, and consequently

$$\text{if} \qquad t \leq 0, \qquad \delta x^j = \delta x^j_0, \qquad \text{with} \qquad \delta x^j_0 = \quad \text{const.} \tag{11.75}$$

Since $h_{\mu\nu}$ is a small perturbation, we expect that the relative position of the particles will changes only by infinitesimal quantities

$$\text{if} \qquad t > 0, \qquad \delta x^{\lambda}(t) = \delta x^{\lambda}_0 + \delta x^{\lambda}_1(t), \tag{11.76}$$

where $\delta x_1^\lambda(t)$ has to be considered as a small perturbation with respect to the initial position $\delta x_0^\lambda$ . Inserting eq. (11.76) in eq. (11.73), remembering that $\delta x_0^\lambda$ is a constant and retaining only first order terms, eq. (11.73) becomes

$$\frac{d^2}{dt^2}\delta x_1^\lambda = \frac{1}{2}\frac{\partial^2 h^{TT\lambda}{}_k}{\partial t^2}\delta x_0^k. \tag{11.77}$$

This equation can be integrated, and the final solution for $\delta x^j$ is

$$\delta x^j = \delta x_0^j + \frac{1}{2}\delta x_0^k h^{TTj}{}_k. \tag{11.78}$$

From eq. (11.78) we see that if the two particles at $t = 0$ lay along the direction of propagation of the wave (the x-axis) and therefore only $\delta x_0^1 \neq 0$, since $h^{TTj}{}_1 = 0$ it follows that their relative position will not be modified by the passage of the wave, as we already established in eq. (11.74).

Let us now study the effect of the polarization of the wave. Consider a plane wave whose nonvanishing components are

$$\begin{aligned} h_{yy} &= -h_{zz} = \Re\left\{A_+ e^{i\omega(t-\frac{x}{c})}\right\} \\ h_{yz} &= h_{zy} = \Re\left\{A_\times e^{i\omega(t-\frac{x}{c})}\right\} \end{aligned} \tag{11.79}$$

Consider two particles initially at a relative distance $y_0$ and $z_0$ from the origin of the frame as indicated in figure (11.1). Let us assume that

$$A_+ \neq 0 \qquad \text{and} \qquad A_\times = 0. \tag{11.80}$$

For example, if $A_+$ is real

$$\Re\left\{A_+ e^{i\omega(t-\frac{x}{c})}\right\} = 2A_+ \cos\omega(t - \frac{x}{c}), \tag{11.81}$$

therefore

$$h_{yy} = -h_{zz} = 2A_+ \cos\omega(t - \frac{x}{c}), \qquad h_{yz} = h_{zy} = 0. \tag{11.82}$$

Assume that the instant $t = 0$ corresponds to $\omega(t - \frac{x}{c}) = \frac{\pi}{2}$ . Remembering that $\delta x^j$ is the distance of the particles 1 and 2 from the origin, and comparing with eq. (11.78) we find that the motion of 1 and 2 for $t > 0$ follows from these equations

$$\begin{array}{lll} 1) & z = 0, & y = y_0 + \frac{1}{2} y_0 h^y{}_y = y0[1 + A_+ \cos\omega(t - \frac{x}{c})], \\ 2) & y = 0, & z = z_0 + \frac{1}{2} z_0 h^z{}_z = z0[1 - A_+ \cos\omega(t - \frac{x}{c})]. \end{array} \tag{11.83}$$

After a quarter of a period ( $\cos\omega(t - \frac{x}{c}) = -1$)

$$\begin{array}{lll} 1) & z = 0, & y = y0[1 - A_+], \\ 2) & y = 0, & z = z0[1 + A_+]. \end{array} \tag{11.84}$$

After half a period ( $\cos\omega(t - \frac{x}{c}) = 0$)

$$\begin{array}{lll} 1) & z = 0, & y = y0, \\ 2) & y = 0, & z = z0. \end{array} \tag{11.85}$$

After three quarters of a period ( $\cos\omega(t - \frac{x}{c}) = 1$)

$$\begin{array}{lll} 1) & z = 0, & y = y0[1 + A_+], \\ 2) & y = 0, & z = z0[1 - A_+]. \end{array} \tag{11.86}$$

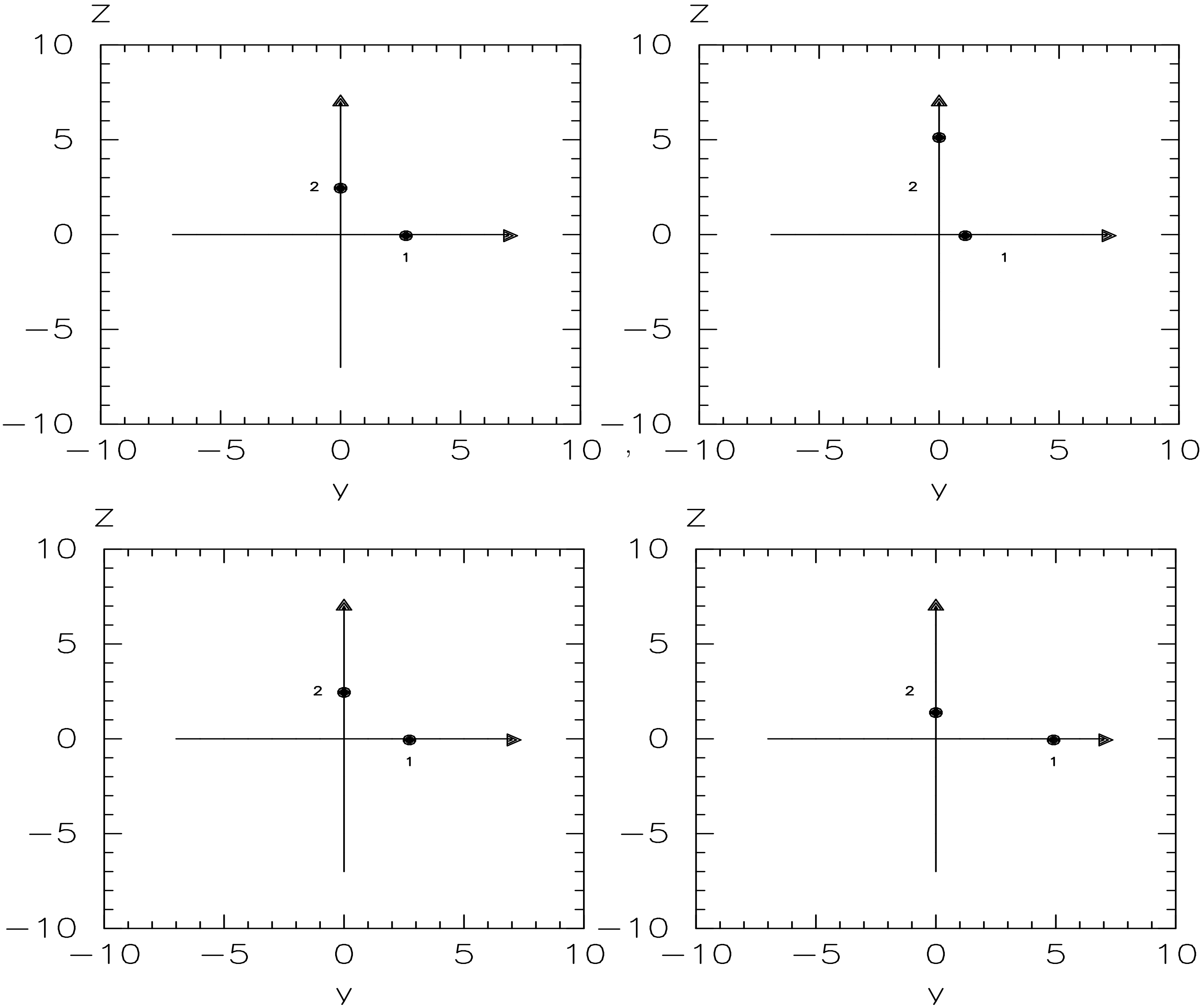



Figure 11.1:

It is now clear that a small ring of particles centered at the origin, will suffer some deformation transforming into a prolate and oblate ellipse in the way indicated in figure (11.2)

Let us now see what happens if $A_\times \neq 0$ and $A_+ = 0$ :

$$h_{yy} = h_{zz} = 0, \qquad h_{yz} = h_{zy} = 2A_\times \cos\omega(t - \frac{x}{c}). \tag{11.87}$$

Comparing with eq. (11.78) we see that a generic particle initially at $P = (y_0, z_0)$, when $t > 0$ will move according to the equations

$$\begin{aligned} y &= y_0 + \frac{1}{2} z_0 h^y{}_z = y_0 + z_0 A_\times \cos\omega(t - \frac{x}{c}), \\ z &= z_0 + \frac{1}{2} y_0 h^z{}_y = z_0 + y_0 A_\times \cos\omega(t - \frac{x}{c}). \end{aligned} \tag{11.88}$$

Let us consider four particles disposed as indicated in figure (11.3) and that $|y_0| = |z_0| = r$. Suppose that the initial time $t = 0$ corresponds to $\omega(t - \frac{x}{c}) = \frac{\pi}{2}$ . The position of the particles will be

$$\begin{array}{llll} 1) & y = r, & z = r, & (11.89) \\ 2) & y = -r, & z = r, & \\ 3) & y = -r, & z = -r, & \\ 4) & y = r, & z = -r, & \end{array}$$

After a quarter of a period $(\cos\omega(t - \frac{x}{c}) = -1)$, the particles will have the following positions

$$1) \qquad y = r[1 - A_\times], \qquad z = r[1 - A_\times], \tag{11.90}$$

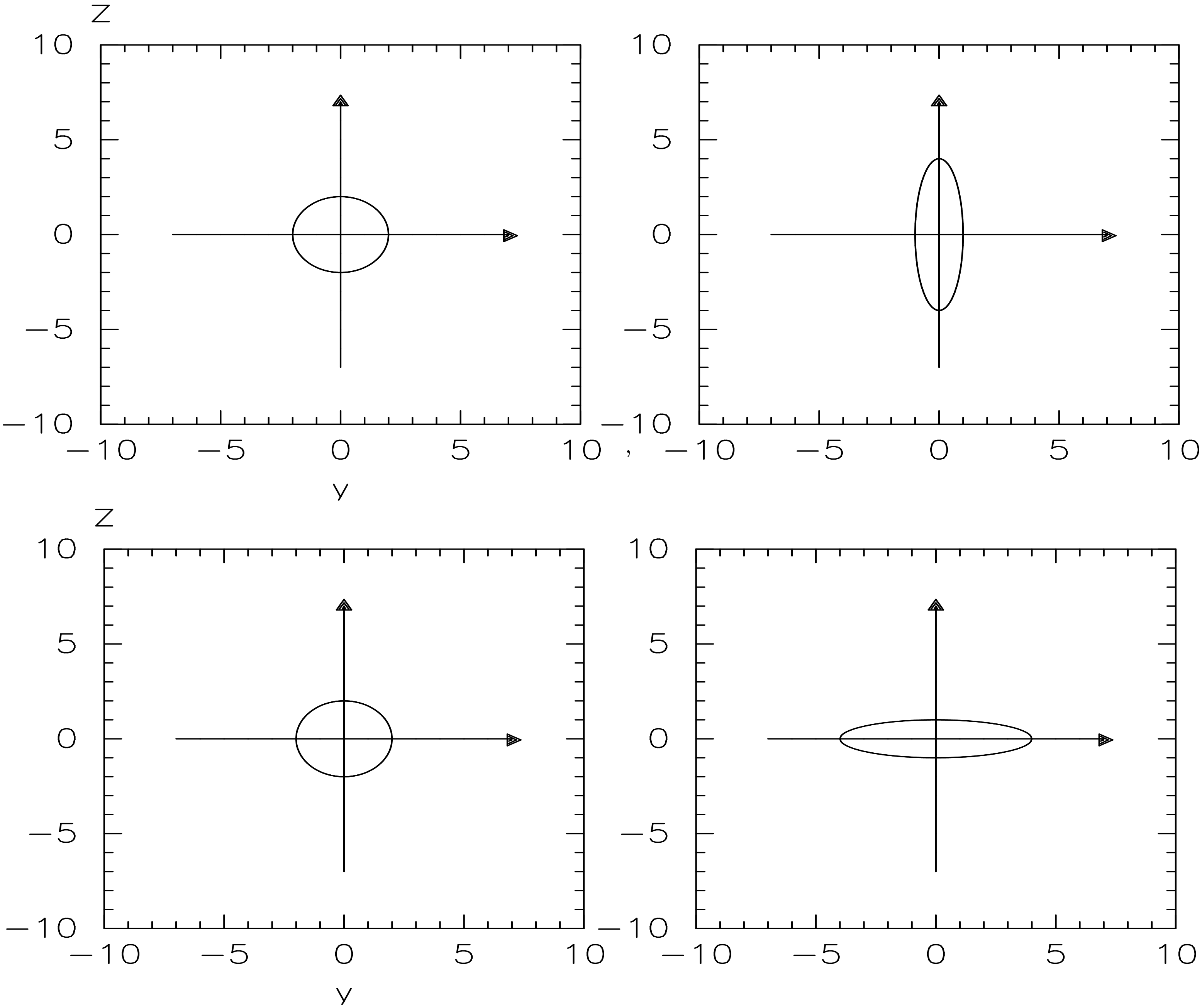



Figure 11.2:

$$
\begin{array}{lll}
2) & y = r[-1 - A_\times], & z = r[1 + A_\times], \\
3) & y = r[-1 + A_\times], & z = r[-1 + A_\times], \\
4) & y = r[1 + A_\times], & z = r[-1 - A_\times],
\end{array}
$$

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 $\cos\omega(t - \frac{x}{c}) = 1$

$$
\begin{array}{lll}
1) & y = r[1 + A_\times], & z = r[1 + A_\times], \\
2) & y = r[-1 + A_\times], & z = r[1 - A_\times], \\
3) & y = r[-1 - A_\times], & z = r[-1 - A_\times], \\
4) & y = r[1 - A_\times], & z = r[-1 + A_\times],
\end{array}
\tag{11.91}
$$

The motion of the particles is indicated in figure (11.3).

It follows that a small ring of particles centered at the origin, will again become an ellipse, but rotated at $45^0$ (see fig. 11.4) 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.

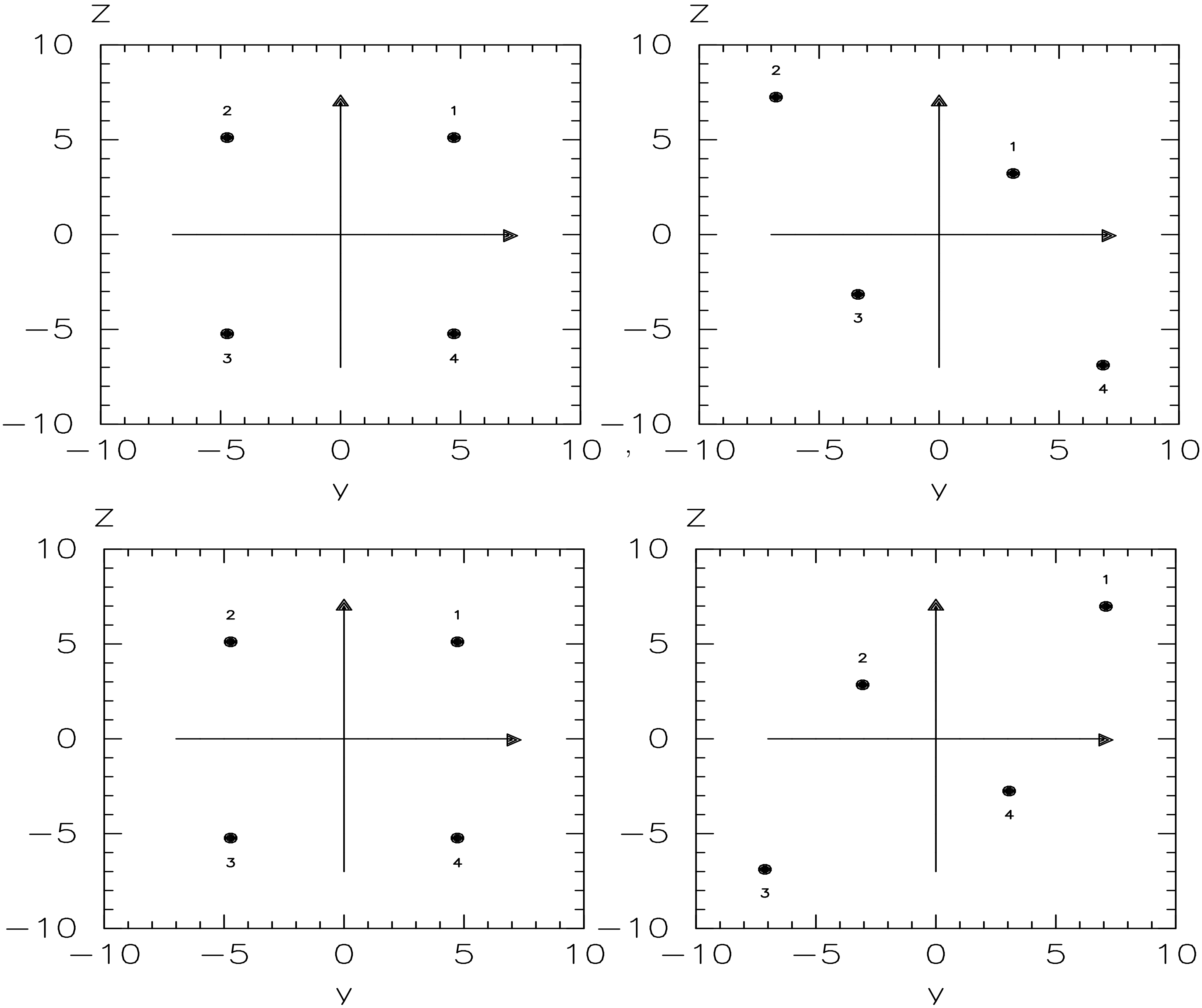



Figure 11.3:

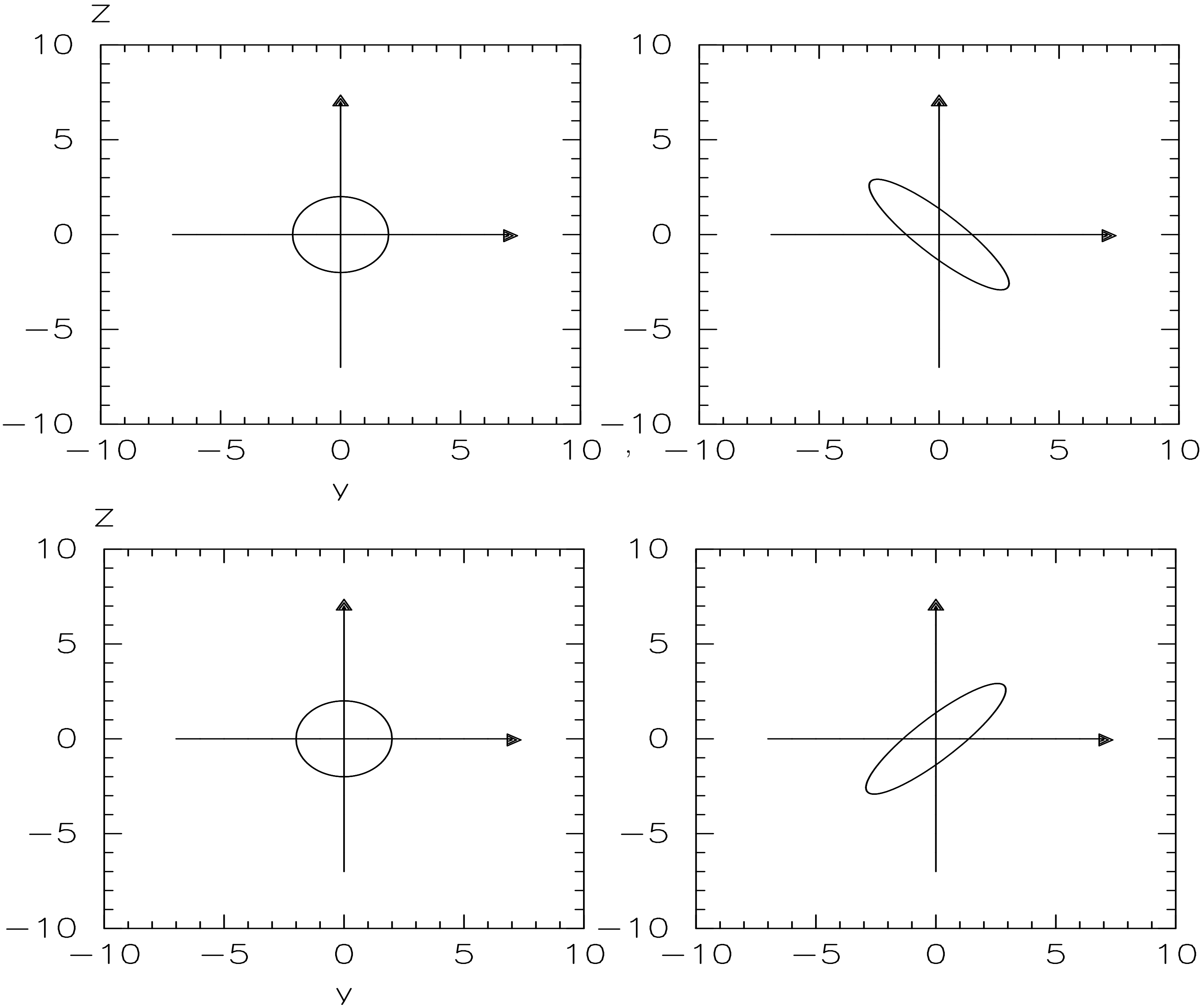



Figure 11.4:

## 11.7 Appendix A

We want to 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

$$x'^{\mu} = x^{\mu} + \epsilon^{\mu} \tag{11.92}$$

provided

$$\Box_F \epsilon_\rho = \frac{\partial h^\beta_\rho}{\partial x^\beta} - \frac{1}{2}\frac{\partial h^\beta_\beta}{\partial x^\rho}. \tag{11.93}$$

When we change the coordinate system the $\Gamma^\lambda$'s transform according to the equation

$$\Gamma'^{\lambda} = \frac{\partial x'^{\lambda}}{\partial x^\rho}\Gamma^\rho - g^{\rho\sigma}\frac{\partial^2 x'^{\lambda}}{\partial x^\rho \partial x^\sigma}, \tag{11.94}$$

where

$$\frac{\partial x'^{\lambda}}{\partial x^\rho} = \frac{\partial \epsilon^\lambda}{\partial x^\rho} + \delta^\lambda_\rho.$$

If $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ (see footnote after eq. (11.14))

$$\Gamma^\rho = g^{\mu\nu}\Gamma^\rho{}_{\mu\nu} = \eta^{\rho k}\left\{ h^\mu{}_{k,\mu} - \frac{1}{2}h^\nu{}_{\nu,k}\right\}. \tag{11.95}$$

Moreover

$$\begin{aligned} g^{\rho\sigma}\frac{\partial^2 x'^{\lambda}}{\partial x^\rho \partial x^\sigma} &= g^{\rho\sigma}\left[\frac{\partial}{\partial x^\rho}\left(\frac{\partial x^\lambda}{\partial x^\sigma} + \frac{\partial \epsilon^\lambda}{\partial x^\sigma}\right)\right] = \\ & g^{\rho\sigma}\left[\frac{\partial}{\partial x^\rho}\left(\delta^\lambda_\sigma + \frac{\partial \epsilon^\lambda}{\partial x^\sigma}\right)\right] \simeq \eta^{\rho\sigma}\left[\frac{\partial^2 \epsilon^\lambda}{\partial x^\rho \partial x^\sigma}\right] = \Box_F \epsilon^\lambda, \end{aligned} \tag{11.96}$$

therefore in the new gauge the condition $\Gamma'^{\lambda} = 0$ becomes

$$\Gamma'^{\lambda} = \eta^{\rho k} \left[ \delta^{\lambda}_{\rho} + \frac{\partial \epsilon^{\lambda}}{\partial x^{\rho}} \right] \left[ \frac{\partial h^{\mu}_{k}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial h^{\nu}{}_{\nu}}{\partial x^{k}} \right] - \Box_F \epsilon^{\lambda} = 0. \tag{11.97}$$

If we neglect second order terms eq.(11.97) becomes

$$\Gamma'^{\lambda} = \eta^{\lambda k} \left[ \frac{\partial h^{\mu}_{k}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial h^{\nu}{}_{\nu}}{\partial x^{k}} \right] - \Box_F \epsilon^{\lambda} = 0.$$

Contracting with $\eta_{\lambda\alpha}$ and remembering that $\eta_{\lambda\alpha}\eta^{\lambda k} = \delta^{k}_{\alpha}$ we finally find

$$\Box_F \epsilon_{\alpha} = \frac{\partial h^{\mu}{}_{\alpha}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial h^{\nu}{}_{\nu}}{\partial x^{\alpha}}.$$

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.

# Chapter 12

# The stress-energy pseudotensor

The divergenceless equation satisfied by the stress-energy tensor can be written in the following form

$$T^{\mu}{}_{\nu;\mu} = \frac{1}{\sqrt{-g}}\frac{\partial\left(T^{\mu}{}_{\nu}\sqrt{-g}\right)}{\partial x^{\mu}} - \frac{1}{2}\frac{\partial g_{\mu\gamma}}{\partial x^{\nu}}T^{\mu\gamma} = 0. \tag{12.1}$$

Let us choose a coordinate system such that all first derivatives of the metric are zero (we know we can do it because of the equivalence principle). In this frame eq. (12.1) becomes

$$\frac{\partial T^{\mu}{}_{\nu}}{\partial x^{\mu}} = 0, \tag{12.2}$$

or, in the contravariant form

$$\frac{\partial T^{\mu\nu}}{\partial x^{\mu}} = 0. \tag{12.3}$$

and consequently $T^{\mu\nu}$ can be expressed as

$$T^{\mu\nu} = \frac{\partial}{\partial x^{\gamma}}\eta^{\mu\nu\gamma}, \tag{12.4}$$

where $\eta^{\mu\nu\gamma}$ is antisymmetric in the indices $\mu$ and $\gamma$. In order to find the explicit expression of $\eta^{\mu\nu\gamma}$ we shall start writing the Einstein equations

$$T^{\mu\nu} = \frac{c^4}{8\pi G}\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right) \tag{12.5}$$

and, remembering that the Riemann tensor is

$$\begin{aligned} R_{\mu\nu\gamma\sigma} = \frac{1}{2}\left(\frac{\partial^2 g_{\mu\sigma}}{\partial x^\nu \partial x^\gamma} + \frac{\partial^2 g_{\nu\gamma}}{\partial x^\mu \partial x^\sigma} - \frac{\partial^2 g_{\mu\gamma}}{\partial x^\nu \partial x^\sigma} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^\mu \partial x^\gamma}\right) \\ + g_{\alpha\beta}\left(\Gamma^\alpha_{\nu\gamma}\Gamma^\beta_{\mu\sigma} - \Gamma^\alpha_{\nu\sigma}\Gamma^\beta_{\mu\gamma}\right) \end{aligned} \tag{12.6}$$

it follows that in the locally inertial frame

$$R^{\mu\nu} = \frac{1}{2}g^{\mu\alpha}g^{\nu\beta}g^{\gamma\delta}\left(\frac{\partial^2 g_{\gamma\beta}}{\partial x^\alpha \partial x^\delta} + \frac{\partial^2 g_{\alpha\delta}}{\partial x^\gamma \partial x^\beta} - \frac{\partial^2 g_{\gamma\delta}}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2 g_{\alpha\beta}}{\partial x^\gamma \partial x^\delta}\right), \tag{12.7}$$

By using this expression, eq. (12.5) becomes

$$T^{\mu\nu} = \frac{\partial}{\partial x^\alpha}\left\{\frac{c^4}{16\pi G}\frac{1}{(-g)}\frac{\partial}{\partial x^\beta}\left[(-g)\left(g^{\mu\nu}g^{\alpha\beta} - g^{\mu\alpha}g^{\nu\beta}\right)\right]\right\} \tag{12.8}$$

The expression in parentheses is antisymmetric in the indices $\nu$ and $\alpha$ and it is the quantity $\eta^{\mu\nu\gamma}$ we were looking for. Since $g_{\mu\nu,\alpha} = 0$ we can extract the factor $\frac{1}{(-g)}$ and write eq. (12.8) as

$$\frac{\partial \zeta^{\mu\nu\alpha}}{\partial x^\alpha} = (-g)T^{\mu\nu}, \tag{12.9}$$

where

$$\zeta^{\mu\nu\alpha} = \frac{c^4}{16\pi G}\frac{\partial}{\partial x^\beta}\left[(-g)\left(g^{\mu\nu}g^{\alpha\beta} - g^{\mu\alpha}g^{\nu\beta}\right)\right]. \tag{12.10}$$

It should be reminded that eq. (12.9) has been derived under the assumption that all first derivatives of the metric tensor vanish. In any other reference the difference $\frac{\partial \zeta^{\mu\nu\alpha}}{\partial x^{\alpha}} - (-g)T^{\mu\nu}$ is not zero. Let us call the difference $(-g)t^{\mu\nu}$ i.e.

$$(-g)\left(T^{\mu\nu} + t^{\mu\nu}\right) = \frac{\partial \zeta^{\mu\nu\alpha}}{\partial x^{\alpha}}. \tag{12.11}$$

The quantities $t^{\mu\nu}$ are symmetric because both $T^{\mu\nu}$ and $\frac{\partial \zeta^{\mu\nu\alpha}}{\partial x^{\alpha}}$ are symmetric in $\mu$ and $\nu$. By expressing $T^{\mu\nu}$ in terms of Einstein's equations and by using eq. (12.10) it is possible to find, after some careful elaboration of the equations

$$\begin{aligned}
t^{\mu\nu} &= \frac{c^4}{16\pi G}\Big\{\left(2\Gamma^{\delta}{}_{\alpha\beta}\Gamma^{\sigma}{}_{\delta\sigma} - \Gamma^{\delta}{}_{\alpha\sigma}\Gamma^{\sigma}{}_{\beta\delta} - \Gamma^{\delta}{}_{\alpha\delta}\Gamma^{\sigma}{}_{\beta\sigma}\right)\left(g^{\mu\alpha}g^{\nu\beta} - g^{\mu\nu}g^{\alpha\beta}\right) \\
&+ g^{\mu\alpha}g^{\beta\delta}\left(\Gamma^{\nu}{}_{\alpha\sigma}\Gamma^{\sigma}{}_{\beta\delta} + \Gamma^{\nu}{}_{\beta\delta}\Gamma^{\sigma}{}_{\alpha\sigma} - \Gamma^{\nu}{}_{\delta\sigma}\Gamma^{\sigma}{}_{\alpha\beta} - \Gamma^{\nu}{}_{\alpha\beta}\Gamma^{\sigma}{}_{\delta\sigma}\right) \\
&+ g^{\nu\alpha}g^{\beta\delta}\left(\Gamma^{\mu}{}_{\alpha\sigma}\Gamma^{\sigma}{}_{\beta\delta} + \Gamma^{\mu}{}_{\beta\delta}\Gamma^{\sigma}{}_{\alpha\sigma} - \Gamma^{\mu}{}_{\delta\sigma}\Gamma^{\sigma}{}_{\alpha\beta} - \Gamma^{\mu}{}_{\alpha\beta}\Gamma^{\sigma}{}_{\delta\sigma}\right) \\
&+ g^{\alpha\beta}g^{\delta\sigma}\left(\Gamma^{\mu}{}_{\alpha\delta}\Gamma^{\nu}{}_{\beta\sigma} - \Gamma^{\mu}{}_{\alpha\beta}\Gamma^{\nu}{}_{\delta\sigma}\right)\Big\}
\end{aligned}$$

Let us consider for example the metric of a plane gravitational wave propagating along the $x$-axis, in the TT-gauge. Let us assume that the metric has only one polarization

$$h_{\mu\nu} = \begin{pmatrix}
(t) & (x) & (y) & (z) \\
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -[1 + h_+(u)] & 0 \\
0 & 0 & 0 & -[1 - h_+(u)]
\end{pmatrix}, \tag{12.12}$$

where $u = t - x$, and $h(u) << 1$. The non-vanishing Christoffel symbols are

$$
\begin{aligned}
\Gamma^{y}{}_{ty} &= -\Gamma^{y}{}_{xy} = \frac{1}{2}\frac{\dot{h}}{(1+h)} \sim \frac{1}{2}\dot{h}, \\
\Gamma^{z}{}_{tz} &= -\Gamma^{z}{}_{xz} = \frac{1}{2}\frac{\dot{h}}{(-1+h)} \sim -\frac{1}{2}\dot{h}, \\
\Gamma^{t}{}_{yy} &= \Gamma^{x}{}_{yy} = \frac{1}{2}\dot{h}, \qquad \Gamma^{t}{}_{zz} = \Gamma^{x}{}_{zz} = -\frac{1}{2}\dot{h}.
\end{aligned}
\tag{12.13}
$$

It is then easy to show that in this case the energy flow across a surface perpendicular to the direction of propagation is

$$
t^{01} = \frac{c^4}{16\pi G}\left[\dot{h}(u)^2\right]. \tag{12.14}
$$

In general, if both polarization are present and

$$
h_{\mu\nu} = \begin{array}{c}
\begin{array}{cccc} (t) & (x) & (y) & (z) \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -[1+h_{+}(u)] & h_{\times}(u) \\
0 & 0 & h_{\times}(u) & -[1-h_{+}(u)]
\end{pmatrix}
\end{array}, \tag{12.15}
$$

$$
t^{01} = \frac{c^4}{16\pi G}\left[\dot{h}_{+}(u)^2 + \dot{h}_{\times}(u)^2\right]. \tag{12.16}
$$

# Chapter 13

# 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. Consider for example the newtonian equations of 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 of constant radius, and the equations of newtonian gravity are symmetric with respect to time translations

$$t' \to t + \tau.$$

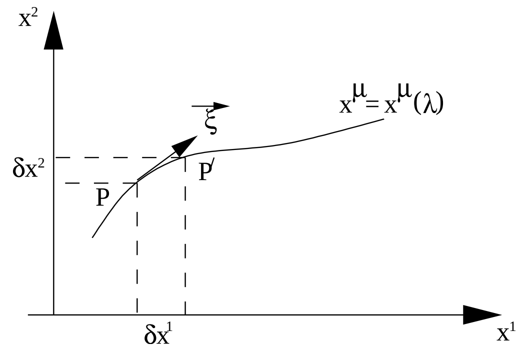
Thus a symmetry corresponds to invariance under translations along certain lines or over certain surfaces. This definition can be applied and extended to Riemannian geometry. We say that a solution of Einstein's equations has a symmetry if there exists an n-dimensional manifold of points with $1 \leq n \leq 4$ , 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 we may have more complicated four-dimensional symmetries. These definitions can be made more precise by introducing the notion of Killing vectors.

## 13.1 The Killing vectors

Consider a vector field $\vec{\xi}(x^\mu)$ defined at every point $x^\alpha$ of a region of spacetime. $\vec{\xi}$ identifies a symmetry if an infinitesimal translation along $\vec{\xi}$ leaves the metric unchanged. $\vec{\xi}$ is the tangent vector to some curve $x^\alpha(\lambda)$ , i.e. $\xi^\alpha = \frac{\delta x^\alpha}{d\lambda}$ , therefore an infinitesimal translation in the direction of $\vec{\xi}$ is an infinitesimal translation along the curve

$$x^{\mu'} = x^\mu + \delta x^\mu. \tag{13.1}$$

Let us consider the 2-dimensional space indicated in the following figure




$$P = (x^1, x^2)$$

$$P' = (x^1 + \delta x^1, x^2 + \delta x^2)$$

Since

$$\delta x^1 = \frac{dx^1}{d\lambda} d\lambda = \xi^1 d\lambda \quad \text{and} \quad \delta x^2 = \frac{dx^2}{d\lambda} d\lambda = \xi^2 d\lambda \tag{13.2}$$

eq. (13.1) becomes

$$x^{\mu'} = x^\mu + \xi^\mu d\lambda. \tag{13.3}$$

When we go from $P$ to $P'$ the metric changes as follows

$$\begin{aligned} g_{\alpha\beta}(P') &\simeq g_{\alpha\beta}(P) + \frac{\partial g_{\alpha\beta}}{\partial \lambda} d\lambda + ... \\ &= g_{\alpha\beta}(P) + \frac{\partial g_{\alpha\beta}}{\partial x^\mu} \frac{\delta x^\mu}{d\lambda} d\lambda + ... \\ &= g_{\alpha\beta}(P) + g_{\alpha\beta,\mu} \xi^\mu d\lambda, \end{aligned} \tag{13.4}$$

hence

$$\delta g_{\alpha\beta} = g_{\alpha\beta,\mu} \xi^\mu d\lambda. \tag{13.5}$$

If there is a symmetry associated to a translation along $\vec{\xi}$ the line-element should not change

$$\delta(ds^2) = \delta(g_{\alpha\beta} \delta x^\alpha \delta x^\beta) = 0, \tag{13.6}$$

hence

$$\delta g_{\alpha\beta}\delta x^{\alpha}\delta x^{\beta}+g_{\alpha\beta}\left[\delta(\delta x^{\alpha})\delta x^{\beta}+\delta x^{\alpha}\delta(\delta x^{\beta})\right]=0. \tag{13.7}$$

Since

$$\begin{aligned}\delta(\delta x^{\alpha}) &= d(\delta x^{\alpha})=d(\xi^{\alpha}d\lambda)=d\xi^{\alpha}d\lambda \\ &= \frac{\partial\xi^{\alpha}}{\partial x^{\mu}}\delta x^{\mu}d\lambda=\xi^{\alpha}_{,\mu}\delta x^{\mu}d\lambda,\end{aligned} \tag{13.8}$$

using eqs. (13.8) and (13.5) eq. (13.7) becomes

$$g_{\alpha\beta,\mu}\xi^{\mu}d\lambda\delta x^{\alpha}\delta x^{\beta}+g_{\alpha\beta}\left[\xi^{\alpha}_{,\mu}\delta x^{\mu}d\lambda\delta x^{\beta}+\xi^{\beta}_{,\gamma}\delta x^{\gamma}d\lambda\delta x^{\alpha}\right]=0. \tag{13.9}$$

After relabelling the indices it becomes

$$\left[g_{\alpha\beta,\mu}\xi^{\mu}+g_{\delta\beta}\xi^{\delta}_{,\alpha}+g_{\alpha\delta}\xi^{\delta}_{,\beta}\right]\delta x^{\alpha}\delta x^{\beta}d\lambda=0. \tag{13.10}$$

Finally, a solution of Einstein's equations will be invariant under translations along $\vec{\xi}$ if and only if

$$g_{\alpha\beta,\mu}\xi^{\mu}+g_{\delta\beta}\xi^{\delta}_{,\alpha}+g_{\alpha\delta}\xi^{\delta}_{,\beta}=0 \tag{13.11}$$

In order to find the Killing vectors of a given a metric $g_{\alpha\beta}$ we have to solve eq. (13.11), which is a system of differential equations for the components of $\vec{\xi}(x^{\mu})$ . If eq. (13.11) admits no solution, the spacetime has no symmetries. It may look like eq. (13.11) is not covariant, since it contains partial derivatives, but it is easy to show that it is equivalent to the following equation (see appendix A)

$$\xi_{\alpha;\beta}+\xi_{\beta;\alpha}=0. \tag{13.12}$$

This is called the **Killing equation**.

The variation of a tensor in the direction of a tensor field $\vec{\xi}$ is called **Lie-derivative** ( $\vec{\xi}$ must not necessarily be a Killing vector), and it is indicated as $L_{\vec{\xi}}$ .

$$L_{\vec{\xi}} T_{\alpha\beta} = T_{\alpha\beta,\mu}\xi^{\mu} + T_{\delta\beta}\xi^{\delta}_{,\alpha} + T_{\alpha\delta}\xi^{\delta}_{,\beta}. \tag{13.13}$$

For the metric tensor

$$L_{\vec{\xi}} g_{\alpha\beta} = g_{\alpha\beta,\mu}\xi^{\mu} + g_{\delta\beta}\xi^{\delta}_{,\alpha} + g_{\alpha\delta}\xi^{\delta}_{,\beta} = \xi_{\alpha;\beta} + \xi_{\beta;\alpha}, \tag{13.14}$$

and if $\vec{\xi}$ is a Killing vector the Lie-derivative of $g_{\alpha\beta}$ must vanish.

The existence of Killing vectors remarkably simplifies the problem of choosing the coordinate system. For example if $\vec{\xi}$ is a timelike vector, we may choose the time axis aligned with $\vec{\xi}$ in such a way that the time coordinate line coincides with the worldline to which $\vec{\xi}$ is tangent and consequently

$$\xi^{\alpha} = (\xi^{0}, 0, 0, 0). \tag{13.15}$$

If, in addition, $\vec{\xi}$ is a Killing vector from eq. (13.11) it follows that

$$\frac{\partial g_{\alpha\beta}}{\partial x^{0}} = 0. \tag{13.16}$$

This means that **if the metric admits a timelike Killing vector, it is independent on time** (A similar procedure can be used if the metric admits a spacelike, or a null killing vector). The map

$$f_t : M \to M$$

under which the metric is unchanged is called an *isometry*, and the Killing vector field is the generator of the isometry.

$\vec{\xi}$ is the tangent vector to some curve $x^\alpha(\lambda)$ , i.e. $\xi^\alpha = \frac{\delta x^\alpha}{d\lambda}$ ,

Given a vector field $\vec{\xi}$ we can find the corresponding congruence of worldlines by integrating the equations

$$\frac{\delta x^\mu}{d\lambda} = \xi^\mu(x^j). \tag{13.17}$$

## 13.2 Examples

*1) The Killing vectors of flat spacetime*

The Killing vectors of Minkowski spacetime can be obtained very easily in cartesian coordinates. Since all Christoffel symbols vanish, the Killing equation becomes

$$\xi_{\alpha,\beta} + \xi_{\beta,\alpha} = 0. \tag{13.18}$$

If one combines the following equations

$$\xi_{\alpha,\beta\gamma} + \xi_{\beta,\alpha\gamma} = 0, \qquad \xi_{\beta,\gamma\alpha} + \xi_{\gamma,\beta\alpha} = 0, \qquad \xi_{\gamma,\alpha\beta} + \xi_{\alpha,\gamma\beta} = 0, \tag{13.19}$$

by using eq. (13.18) we find

$$\xi_{\alpha,\beta\gamma} = 0, \tag{13.20}$$

whose general solution is

$$\xi_\alpha = c_\alpha + \epsilon_{\alpha\gamma} x^\gamma. \tag{13.21}$$

Substitution into eq. (13.18) gives

$$\epsilon_{\alpha\gamma}x^{\gamma}_{,\beta} + \epsilon_{\beta\gamma}x^{\gamma}_{,\alpha} = \epsilon_{\alpha\gamma}\delta^{\gamma}_{\beta} + \epsilon_{\beta\gamma}\delta^{\gamma}_{\alpha} = \epsilon_{\alpha\beta} + \epsilon_{\beta\alpha} = 0$$

Therefore eq. (13.21) is the solution only if

$$\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha} \tag{13.22}$$

Thus flat spacetime possesses ten linearly independent Killing vectors (the 4 constants $c_{\alpha}$ and the six independent components of $\epsilon_{\alpha\beta}$ ).

*2) The Killing vectors of a spherical surface*

Consider a sphere of unit radius

$$ds^2 = d\theta^2 + sin^2\theta d\varphi^2 = (dx^1)^2 + sin^2 x^1 (dx^2)^2. \tag{13.23}$$

The Killing equation written in the form of eq. (13.11)

$$g_{\alpha\beta,\mu}\xi^{\mu} + g_{\delta\beta}\xi^{\delta}_{,\alpha} + g_{\alpha\delta}\xi^{\delta}_{,\beta} = 0$$

gives

$$1) \quad \alpha = \beta = 1 \quad 2g_{\delta 1}\xi^{\delta}_{,1} = 0 \rightarrow \xi^{1}_{,1} = 0 \tag{13.24}$$

$$2) \quad \alpha = 1, \beta = 2 \quad g_{\delta 2}\xi^{\delta}_{,1} + g_{1\delta}\xi^{\delta}_{,2} = 0 \rightarrow \xi^{1}_{,2} + \sin^2\theta\xi^{2}_{,1} = 0$$

$$3) \quad \alpha = \beta = 2 \quad g_{22,\mu}\xi^{\mu} + 2g_{\delta 2}\xi^{\delta}_{,2} = 0 \rightarrow \cos\theta\xi^{1} + \sin\theta\xi^{2}_{,2} = 0.$$

The general solution is

$$\xi^1 = Asin(\varphi + a), \qquad \xi^2 = Acos(\varphi + a)cot\theta + b. \tag{13.25}$$

Therefore a spherical surface admits three linearly independent Killing vectors depending on the choice of the integration constants $(A, a, b)$.

## 13.3 Conserved quantities in geodesic motion

In Chapter 6, section 2, we showed that geodesics are those curves which parallel-transport their own tangent vector.

$$\nabla_{\vec{U}}\vec{U} = 0, \tag{13.26}$$

or

$$\frac{dU^\alpha}{d\tau} + \Gamma^\alpha{}_{\beta\nu}U^\beta U^\nu = 0. \tag{13.27}$$

(see eqs. (6.11) and (6.12)). If we consider a particle moving along a geodesic, and choose the affine parameter as being the proper time, the tangent vector will be the four-velocity of the particle $\vec{U} = \frac{\delta x^\alpha}{d\tau}$. Let us assume that the metric admits a Killing vector $\vec{\xi}$. If we contract eq. (13.27) with $\vec{\xi}$ we find

$$\xi_\alpha\left[\frac{dU^\alpha}{d\tau} + \Gamma^\alpha{}_{\beta\nu}U^\beta U^\nu\right] = \frac{d(\xi_\alpha U^\alpha)}{d\tau} - U^\alpha\frac{d\xi_\alpha}{d\tau} + \Gamma^\alpha{}_{\beta\nu}U^\beta U^\nu\xi_\alpha. \tag{13.28}$$

Since

$$U^\alpha\frac{d\xi_\alpha}{d\tau} = U^\beta\frac{d\xi_\beta}{d\tau} = U^\beta\frac{\partial\xi_\beta}{\partial x^\nu}\frac{\delta x^\nu}{d\tau} = U^\beta U^\nu\frac{\partial\xi_\beta}{\partial x^\nu}, \tag{13.29}$$

and eq. (13.28) becomes

$$\frac{d(\xi_\alpha U^\alpha)}{d\tau} - U^\beta U^\nu\left[\frac{d\xi_\beta}{dx^\nu} - \Gamma^\alpha{}_{\beta\nu}\xi_\alpha\right] = 0, \tag{13.30}$$

or

$$\frac{d(\xi_\alpha U^\alpha)}{d\tau} - U^\beta U^\nu\xi_{\beta;\nu} = 0. \tag{13.31}$$

Since $\xi_{\beta;\nu}$ is antisymmetric in $\beta$ and $\nu$, when contracted on $U^\beta U^\nu$ which is symmetric, gives zero as a result. Finally eq. (13.31) becomes

$$\frac{d(\xi_\alpha U^\alpha)}{d\tau} = 0. \tag{13.32}$$

This means that the quantity

$$\xi_\alpha U^\alpha = const, \tag{13.33}$$

is a constant of motion for the particle. Thus in mechanics for every Killing vector there is an associated conserved quantity. For example, if $\vec{\xi}$ is a timelike Killing vectors and we choose the coordinates in such a way that $\xi^\alpha = (\xi^0, 0, 0, 0)$, eq. (13.33) becomes

$$\xi_0 U^0 = const, \tag{13.34}$$

which expresses the conservation of the energy of the particle.

However 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.

## 13.4 Killing vectors and conservation laws

In the chapter on the stress-energy tensor, we have shown that it satisfies the "conservation law"

$$T^{\mu\nu}{}_{;\nu} = 0, \tag{13.35}$$

and we have shown that in general this is not a genuine conservation law. If the spacetime admits a Killing vector, then

$$\left(\xi_\mu T^{\mu\nu}\right)_{;\nu} = \xi_{\mu;\nu} T^{\mu\nu} + \xi_\mu T^{\mu\nu}{}_{;\nu} = 0. \tag{13.36}$$

In fact the second term vanishes because of eq. (13.35) and the first vanishes because $\xi_{\mu;\nu}$ is antisymmetric, and $T^{\mu\nu}$ is symmetric. But now the quantity $(\xi_\mu T^{\mu\nu})$ is a vector, and according to eq. (8.57)

$$V^\nu{}_{;\nu} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\nu} \left(\sqrt{-g} V^\nu\right), \tag{13.37}$$

therefore eq. (13.36) is equivalent to

$$\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\nu} \left[\sqrt{-g} \left(\xi_\mu T^{\mu\nu}\right)\right] = 0, \tag{13.38}$$

and accordingly, a conserved quantity can be defined as

$$T = \int_{(x^0 = const)} \sqrt{-g} \left(\xi_\mu T^{\mu 0}\right) dx^1 dx^2 dx^3, \tag{13.39}$$

as we showed in the chapter on the stress-energy tensor. (For a more detailed discussion of Gauss' theorem see the next chapter). If the Killing vector is timelike , the associated conserved quantity will be called energy. (Remember that in classical mechanics energy is conserved when the hamiltonian is time independent, thus conservation of energy is associated to a symmetry with respect to translations in time, which is precisely what the existence of a timelike Killing vector means, except that now 'time' means the $x^0$-coordinate). Thus the energy of a gravitational system can be defined in a non ambiguous way only if there exists a timelike Killing vector field. Similarly when there is a spacelike Killing vector the associated conserved quantities are called momentum or angular momentum, though this is more a matter of definition.

## 13.5 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 are perpendicular to that surface, we say that $\vec{V}$ is **hypersuface orthogonal**. If this is the case, $\vec{V}$ should be parallel to the gradient of the family of surfaces. Let us clarify what do we mean. As we described in chapter 2, page 43, the gradient of a surface $f(x^\mu)$ is a one-form

$$\tilde{d}f \to (\frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, ... \frac{\partial f}{\partial x^n}) = \{f_{,\alpha}\}. \tag{13.40}$$

When we say that $\vec{V}$ must be parallel to $\tilde{d}f$ we mean that the one-form dual to $\vec{V}$ , i.e. $\tilde{V} \to \{g_{\alpha\beta}V^\beta \equiv V_\alpha\}$ must satisfy the equation

$$\frac{\partial f}{\partial x^\alpha} = \lambda V_\alpha, \qquad \text{or} \qquad V_\alpha = \frac{1}{\lambda} f_{,\alpha}. \tag{13.41}$$

If this is true, it follows that

$$\begin{aligned} V_{\alpha;\beta} - V_{\beta;\alpha} &= \left(\frac{f_{,\alpha}}{\lambda}\right)_{;\beta} - \left(\frac{f_{,\beta}}{\lambda}\right)_{;\alpha} \\ &= \frac{1}{\lambda}\left(f_{,\alpha;\beta} - f_{,\beta;\alpha}\right) + f_{,\alpha}(\lambda^{-1})_{;\beta} - f_{,\beta}(\lambda^{-1})_{;\alpha} = \\ &= \frac{1}{\lambda}\left(f_{,\alpha,\beta} - f_{,\beta,\alpha} - \Gamma^{\mu}{}_{\beta\alpha} f_{,\mu} + \Gamma^{\mu}{}_{\alpha\beta} f_{,\mu}\right) + f_{,\alpha}(-\frac{\lambda_{,\beta}}{\lambda^2}) - f_{,\beta}(-\frac{\lambda_{,\alpha}}{\lambda^2}) \\ &= V_\beta \frac{\lambda_{,\alpha}}{\lambda} - V_\alpha \frac{\lambda_{,\beta}}{\lambda}, \end{aligned} \tag{13.42}$$

i.e.

$$V_{\alpha;\beta} - V_{\beta;\alpha} = V_\beta \frac{\lambda_{,\alpha}}{\lambda} - V_\alpha \frac{\lambda_{,\beta}}{\lambda}. \tag{13.43}$$

If we now define the following quantity

$$\omega^\delta = \frac{1}{2}\epsilon^{\delta\alpha\beta\mu} V_{[\alpha;\beta]} V_\mu, \tag{13.44}$$

which we call *rotation*, remembering the definition of the antisymmetric unit pseudotensor $\epsilon^{\delta\alpha\beta\mu}$ given in Appendix B, one can show that

$$\omega^\delta = 0. \tag{13.45}$$

This is the Frobenius theorem, which states the sufficient and necessary condition for $\vec{V}$ to be a hypersurface-orthogonal vector field.

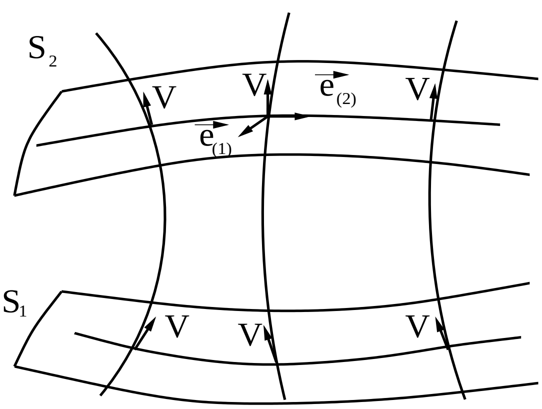
What do we do with hypersurface-orthogonal vectors? In general, if we have any vector field, for example a timelike vector field $V^\alpha(x^\mu)$, we can choose the worldlines of $\vec{V}$ as coordinate lines such that at any point the vector field takes a very simple form

$$V^\alpha(x^\mu) \rightarrow (V^0(x^\mu), 0, 0, 0), \tag{13.46}$$

and by means of a coordinate transformation $x^{0\prime} = x^{0\prime}(x^\mu)$ we can set $V^0 = 1$. (Similarly, if $\vec{V}$ is spacelike we can set $V^\alpha(x^\mu) \to (0, V^1(x^\mu), 0, 0, 0)$ ). If $\vec{V}$ is the four-velocity of some matter, we say that we choose a **comoving** coordinate system. It is always possible to choose coordinates such that $\vec{V}$ takes the simple form (13.46), but the corresponding transformation for the dual one-form $\tilde{V}$

$$V_\alpha(x^\mu) \to (V_0(x^\mu), 0, 0, 0), \tag{13.47}$$

is possible only if $\vec{V}$ is hypersurface-orthogonal. In fact, eq. (13.47) is equivalent to $V_\alpha = V_0 x^0_{,\alpha}$, i.e. if (13.47) holds, $\vec{V}$ is hypersurface-orthogonal and viceversa (remember that $x^\alpha_{,\beta} = \delta^\alpha_\beta$ ). The existence of an hypersurface-orthogonal vector field allows to choose coordinates in such a way that the metric can be considerably simplified. Given a three-dimensional spacetime $(x^0, x^1, x^2)$



be 1,2,3, the worldlines of a vector field $\vec{V}$, and be $S_1$ and $S_2$ the projections of the surfaces to which $\vec{V}$ is orthogonal. We can choose the remaining basis vectors as the tangent vectors to some curves lying on the surface, and $\vec{e}_{(0)}$ parallel to $\vec{V}$, so that

$$g_{00} = g(\vec{e}_{(0)}, \vec{e}_{(0)}) = \vec{e}_{(0)} \cdot \vec{e}_{(0)} \neq 0 \tag{13.48}$$

$$g_{0i} = g(\vec{e}_{(0)}, \vec{e}_{(i)}) = 0, \qquad i = 1, 2,$$

and the metric takes the form

$$ds^2 = g_{00}(dx^0)^2 + g_{ik}(dx^i)(dx^k), \qquad i, k = 1, 3 \tag{13.49}$$

This is possible provided $\vec{V}$ is not a null vector, otherwise $g_{00}$ would be zero. The generalization of this example to the four-dimensional spacetime is immediate, and the surface S will now be a hypersurface.

## 13.6 Appendix A

We want to show that eq. (13.11) is equivalent to eq. (13.12).

$$\begin{aligned} \xi_{\alpha;\beta} &= (g_{\alpha\mu}\xi^{\mu})_{;\beta} \\ &= g_{\alpha\mu}\xi^{\mu}_{;\beta} = g_{\alpha\mu}\left(\xi^{\mu}_{,\beta} + \Gamma^{\mu}{}_{\delta\beta}\xi^{\delta}\right), \end{aligned} \tag{13.50}$$

hence

$$\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = g_{\alpha\mu}\left(\xi^{\mu}_{,\beta} + \Gamma^{\mu}{}_{\delta\beta}\xi^{\delta}\right) \tag{13.51}$$

$$
\begin{aligned}
&+ \quad g_{\beta\mu}\left(\xi^{\mu}_{,\alpha} + \Gamma^{\mu}{}_{\alpha\delta}\xi^{\delta}\right) \\
&= \quad g_{\alpha\mu}\xi^{\mu}_{,\beta} + g_{\beta\mu}\xi^{\mu}_{,\alpha} + \left(g_{\alpha\mu}\Gamma^{\mu}{}_{\delta\beta} + g_{\beta\mu}\Gamma^{\mu}{}_{\alpha\delta}\right)\xi^{\delta}.
\end{aligned}
$$

The term in parenthesis can be written as

$$
\begin{aligned}
&\frac{1}{2}\left[g_{\alpha\mu}g^{\mu\sigma}\left(g_{\delta\sigma,\beta} + g_{\sigma\beta,\delta} - g_{\delta\beta,\sigma}\right) + g_{\beta\mu}g^{\mu\sigma}\left(g_{\alpha\sigma,\delta} + g_{\sigma\delta,\alpha} - g_{\alpha\delta,\sigma}\right)\right] \\
&= \frac{1}{2}\left[\delta^{\sigma}_{\alpha}\left(g_{\delta\sigma,\beta} + g_{\sigma\beta,\delta} - g_{\delta\beta,\sigma}\right) + \delta^{\sigma}_{\beta}\left(g_{\alpha\sigma,\delta} + g_{\sigma\delta,\alpha} - g_{\alpha\delta,\sigma}\right)\right] \qquad (13.52) \\
&= \frac{1}{2}\left[g_{\delta\alpha,\beta} + g_{\alpha\beta,\delta} - g_{\delta\beta,\alpha} + g_{\alpha\beta,\delta} + g_{\beta\delta,\alpha} - g_{\alpha\delta,\beta}\right] \\
&= g_{\alpha\beta,\delta},
\end{aligned}
$$

and eq. (13.51) becomes

$$
\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = g_{\alpha\mu}\xi^{\mu}_{,\beta} + g_{\beta\mu}\xi^{\mu}_{,\alpha} + g_{\alpha\beta,\delta}\xi^{\delta} \qquad (13.53)
$$

which coincides with eq. (13.11).

## 13.7 Appendix B: The Levi-Civita completely antisymmetric pseudotensor

We define $\epsilon^{iklm}$ as an objects 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

$$
\epsilon^{0123} = 1. \qquad (13.54)
$$

Suppose we are in Minkowski spacetime. If we rotate the coordinate system, one can show that $\epsilon^{iklm}$ behaves as a tensor, but if we reverse the sign of one coordinate, the components of $\epsilon^{iklm}$ do not change, since they have been defined as being the same in all coordinate systems, while for a generic tensor some components do change sign. This is the reason why $\epsilon^{iklm}$ is a pseudotensor, in the sense that it behaves as a tensor only for a selected class of coordinate transformations. If all this is true in a locally inertial frame, it must be true in any coordinate system.