Introduction to Astrophysics and Cosmology

Relativistic Cosmology - Elements of GR

Helga Dénes 2024 S1 Yachay Tech

hdenes@yachaytech.edu.ec

GR

Tensors provide a natural mathematical language for describing the curvature of spacetime.

First we will discuss an introduction to tensor analysis and then an introduction to general relativity at a technical level.

It is helpful to clearly distinguish the purely mathematical topics from the physical concepts of general relativity. So, when we develop tensor analysis in the next section, we shall develop it as a purely mathematical subject without bringing in general relativistic concerns at all. The two-dimensional metrics (10.7), (10.8) and (10.9) introduced in §10.2 will be used as illustrative examples repeatedly to clarify various points.

After introducing the basics of tensor analysis in the next section, we shall start developing the basic concepts of general relativity.

There are many excellent textbooks on this subject, to which one can turn in order to learn the subject at a greater depth.

A good book for special relativity: **Robert Resnick: Introduction to Special Relativity**

Tensors

In elementary **physics**, usually **two kinds of physical quantities** are introduced: **scalars and vectors.**

A **vector** is defined as a quantity which has **both magnitude and direction.** There is, however, an alternative way of defining a vector. A vector has components of which each one is associated with one coordinate axis (i.e. components A_x, A_y and A_z of the vector $\mathbf{A}$ are associated with x, y and z axes respectively) and they **transform in some particular way when we change from one coordinate system to another**. Since this alternative definition of vectors provides a natural entry point into the world of tensors, let us consider in some detail how the components of a vector transform on changing the coordinate system.

In order to find out how the components of a vector transform, we first need to know how the components of a **vector are defined in a coordinate system**.

Let us take the concrete **examples of the generalized velocity** and the **generalized force** which we encounter in Lagrangian mechanics.

Tensors

Suppose we denote the generalized coordinates in a system by x^i, where i can have values $i = 1, 2, \ldots, N$.

A component of the **generalized velocity is given by dx^i/dt,** whereas a component of the **generalized force is given by $-\partial V/\partial x^i$** where V **is the potential.**

If components of generalized velocity and generalized force are **defined in another coordinate system** $\bar{x}^i$ in **exactly the same way**, then the **chain rule** of partial differentiation implies

$$\frac{d\bar{x}^i}{dt} = \sum_{k=1}^{N} \frac{dx^k}{dt} \frac{\partial \bar{x}^i}{\partial x^k}, \tag{12.1}$$

$$\frac{\partial V}{\partial \bar{x}^i} = \sum_{k=1}^{N} \frac{\partial V}{\partial x^k} \frac{\partial x^k}{\partial \bar{x}^i}. \tag{12.2}$$

It should be noted that these two transformation laws are slightly different.

Tensors

Vectors **transforming like dx^i/dt are called** ***contravariant vectors*** and are indicated by superscripts, whereas vectors **transforming like $\partial V/\partial x^i$ are called** ***covariant vectors*** and are indicated by subscripts (i in x^i or $\bar{x}^i$ appearing at the bottom of a derivative should be treated as **subscript**).

We also introduce the well-known ***summation convention*** that **if an index is repeated twice in a term, once as a subscript and once as a superscript, then it automatically implies summation over the possible values of that index and it is not necessary to put the summation sign explicitly**.

Using this summation convention, the transformation laws for a contravariant vector A^i and a covariant vector A_i would be

$$\frac{d\bar{x}^i}{dt} = \sum_{k=1}^{N} \frac{dx^k}{dt}\frac{\partial \bar{x}^i}{\partial x^k}, \qquad (12.1)$$

$$\frac{\partial V}{\partial \bar{x}^i} = \sum_{k=1}^{N} \frac{\partial V}{\partial x^k}\frac{\partial x^k}{\partial \bar{x}^i}. \qquad (12.2)$$

$$\overline{A}^i = A^k \frac{\partial \overline{x}^i}{\partial x^k}, \qquad (12.3)$$

$$\overline{A}_i = A_k \frac{\partial x^k}{\partial \overline{x}^i}, \qquad (12.4)$$

where $\bar{A}^i$ and $\bar{A}_i$ are components of these contravariant and covariant vectors in the coordinate system $\bar{x}^i$.

Tensors

By comparing (12.3) with (12.1) and (12.4) with (12.2), it is clear that the generalized **velocity** dx^i/dt transforms as a **contravariant** vector and the generalized **force** $-\partial V/\partial x^i$ transforms as a **covariant** vector.

If we consider the **transformation from one Cartesian frame to another** (for example, due to a rotation from one frame to the other in two dimensions), it is easy to show that

$$\boxed{\frac{\partial \overline{x}^i}{\partial x^k} = \frac{\partial x^k}{\partial \overline{x}^i}}$$

This implies that the **distinction between contravariant and covariant vectors disappears if we consider only transformations between Cartesian frames.**

Tensors

A component of a vector is associated with only one coordinate axis. **In the case of a general tensor, a component can be associated with several coordinate axes**. So a component will generally have several indices. The **transformation law of a general tensor** will be the following:

$$\overline{T}^{ab..d}_{l..n} = T^{\alpha\beta..\delta}_{\lambda..\nu} \frac{\partial \overline{x}^a}{\partial x^\alpha} \frac{\partial \overline{x}^b}{\partial x^\beta} .. \frac{\partial \overline{x}^d}{\partial x^\delta} \frac{\partial x^\lambda}{\partial \overline{x}^l} .. \frac{\partial x^\nu}{\partial \overline{x}^n}. \qquad (12.5)$$

Note that some indices are put as superscripts and some as subscripts depending on whether the corresponding parts of the transformation are like contravariant vectors or covariant vectors.

From the transformation law (12.5) of tensors, it is very easy to show that **the product $A_i B_k$ of two vectors A_i and B_k should transform like a tensor with two covariant components i and k**.

This can be generalized to the result that **the product of two tensors gives a tensor of higher rank**.

Tensors

One very important operation is the ***contraction* of tensors**.

Suppose we write $n = d$ in the tensor $\bar{T}^{ab..d}_{l..n}$

This, by the summation convention, implies that we are summing over all possible values of $n = d$. It follows from (12.5) that

$$\overline{T}^{ab..d}_{l..d} = T^{\alpha\beta..\delta}_{\lambda..\nu} \frac{\partial \overline{x}^a}{\partial x^\alpha} \frac{\partial \overline{x}^b}{\partial x^\beta} .. \frac{\partial \overline{x}^d}{\partial x^\delta} \frac{\partial x^\lambda}{\partial \overline{x}^l} .. \frac{\partial x^\nu}{\partial \overline{x}^d}$$

Using the fact that

$$\frac{\partial \overline{x}^d}{\partial x^\delta} \frac{\partial x^\nu}{\partial \overline{x}^d} = \delta^\nu_\delta$$

we easily get

$$\overline{T}^{ab..d}_{l..d} = T^{\alpha\beta..\delta}_{\lambda..\delta} \frac{\partial \overline{x}^a}{\partial x^\alpha} \frac{\partial \overline{x}^b}{\partial x^\beta} .. \frac{\partial x^\lambda}{\partial \overline{x}^l} .. \qquad (12.6)$$

Tensors

on making use of the obvious result $T^{\alpha\beta..\delta}_{\lambda..\nu}\delta^{\nu}_{\delta} = T^{\alpha\beta..\delta}_{\lambda..\delta}$ following from the properties of the Kronecker δ.

δ^i_k **transforms like a tensor**, but not δ_{ik} or δ^{ik}

(the Kronecker δ in any form is assumed to have the value 1 if $i = k$ and 0 otherwise).

It is clear from (12.6) that $T^{\alpha\beta..\delta}_{\lambda..\delta}$ transforms like a tensor with one contravariant index and one covariant index less compared to $T^{\alpha\beta..\delta}_{\lambda..\nu}$. Thus **the operation of contraction reduces the rank of the tensor** (by reducing one contravariant rank and one covariant rank).

In mathematics, the **Kronecker delta** is a function of two variables, usually just non-negative integers. The function is 1 if the variables are equal, and 0 otherwise:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

The metric tensor

$$ds^2 = \sum_{\alpha,\beta} g_{\alpha\beta}\, dx_\alpha\, dx_\beta, \qquad (10.5)$$

We have pointed out in §10.2 that the distance between two nearby points in a space is given by (10.5). Using our present notation and the summation convention, we can write

$$\boxed{ds^2 = g_{ik}\, dx^i dx^k. \qquad (12.7)}$$

Although x^i does not transform as a vector, ***dx^i* is clearly a contravariant vector**.

Hence $g_{ik} dx^l ds^m$ **must be a tensor** which, **after two contractions, would give a scalar,** implying that ***ds^2* is a scalar.**

We pointed out in §10.2 that it is the **metric tensor** g_{ik} which **determines whether a space is curved or not**.

The metric tensor

We shall now develop the mathematical machinery to **calculate curvature**. As we develop the techniques, we shall illustrate them by applying them to the **metrics**

$$ds^2 = dr^2 + r^2\, d\theta^2, \tag{12.8}$$

$$ds^2 = a^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{12.9}$$

where (12.8) corresponds to a **plane using polar coordinates** and (12.9) corresponds to the **surface of a sphere.**

We shall also sometimes consider the metric

$$ds^2 = a^2(dx_1^2 + \sinh^2 x_1\, dx_2^2), \tag{12.10}$$

which corresponds to the **surface having the property of a saddle point at every point.**

We pointed out in §10.2 that these three metrics can be written in the forms (10.7)–(10.9) by using very similar notation and also discussed that these correspond to the only three possible uniform two-dimensional surfaces (i.e. surfaces in which every point is equivalent).

$$ds^2 = g_{ik}\, dx^i dx^k. \tag{12.7}$$

The metric tensor

It may be noted that (12.7) allows for cross-terms of the form $g_{12}dx^1dx^2$.

The metrics (12.8)–(12.10), however, have only **pure quadratic terms** and no cross-terms. If the coordinate system is orthogonal, then we do not have cross-terms in the expression of ds^2.

We shall **restrict our discussions to only simple metrics** without cross-terms corresponding to orthogonal coordinates. In other words, we shall only be concerned with **metric tensors g_{ik} which have non-zero terms on the diagonal alone when represented in the form of a matrix.**

For the three metrics (12.8)– (12.10), the components of the metric tensor respectively are

$$g_{rr} = 1,\ \ g_{\theta\theta} = r^2,\ \ g_{r\theta} = g_{\theta r} = 0, \tag{12.11}$$

$$g_{\theta\theta} = a^2,\ \ g_{\phi\phi} = a^2 \sin^2\theta,\ \ g_{\theta\phi} = g_{\phi\theta} = 0, \tag{12.12}$$

$$g_{11} = a^2,\ \ g_{22} = a^2 \sinh^2 x_1,\ \ g_{12} = g_{21} = 0. \tag{12.13}$$

There are astrophysical situations where the **cross-terms become important, such as the Kerr metric** of a rotating black hole.

$$\boxed{ds^2 = g_{ik}\, dx^i dx^k. \qquad (12.7)}$$

The metric tensor

From a contravariant vector A^i, it is possible to construct the corresponding covariant vector in the following way

$$A_i = g_{ik} A^k. \qquad (12.14)$$

This is often **called the *lowering of an index*.** It is very **easy to do it if the metric is diagonal.**

For example, if (A^r, A^θ) are the contravariant components of a vector in the plane with the metric tensor (12.11), then the corresponding covariant components are $(A_r = A^r, A_\theta = r^2 A^\theta)$.

Suppose dx^i and dx^k have corresponding covariant vectors dx_i and dx_k obtained according to (12.14). It should be possible to write the metric in the form

$$ds^2 = g^{ik} dx_i dx_k. \qquad (12.15)$$

By requiring that ds^2 given by (12.7) and (12.15) should be equal

$$g^{ik} g_{kl} = \delta^i_l. \qquad (12.16)$$

The metric tensor

For a diagonal metric tensor g_{ik}, it is particularly easy **to obtain the corresponding g^{ik}.**

One merely has to take the **inverse of the diagonal elements** while leaving the off-diagonal elements zero in order to satisfy (12.16).

For example, for the metric tensor (12.12) corresponding to the surface of a sphere, it is easily seen that

$$g^{\theta\theta} = \frac{1}{a^2}, \; g^{\phi\phi} = \frac{1}{a^2 \sin^2\theta}, \; g^{\theta\phi} = g^{\phi\theta} = 0 \tag{12.17}$$

would satisfy (12.16). Once the covariant metric tensor g_{ik} has been introduced, we can use it to raise an index and obtain a contravariant vector from a covariant vector in the following way

$$A^i = g^{ik} A_k. \tag{12.18}$$

The metric tensor

Starting from a **contravariant vector** A^i , if we **once lower the index** by (12.14) and **then raise it** again by (12.18), then it is easy to use (12.16) to show that **we get back the same vector** A^i.

We often have to deal with components of a vector along the coordinate directions. For a particle moving in a plane, for example, the components of velocity in polar coordinates are $(\dot{r}, r\dot{\theta})$. We can call these ***vectorial components***.

It is obvious that the vectorial components in different orthogonal coordinate systems will not transform amongst each other according to either (12.3) or (12.4). However, if we divide the two vectorial components of velocity by $\sqrt{g_{rr}} = 1$ and $\sqrt{g_{\theta\theta}} = r$ respectively, then we get the contravariant velocity vector $(\dot{r}, \dot{\theta})$ in polar coordinates.

In general, if we divide the *i*-th vectorial component of a vector in an orthogonal coordinate system by $\sqrt{g_{ii}}$ and do this for each component, then we get a set of components of the vector which would transform between orthogonal coordinates in accordance with the rule (12.3) for contravariant vectors.

The metric tensor

Similarly, **if we multiply the vectorial components by $\sqrt{g_{ii}}$, then we would get the components of the corresponding covariant vector.**

As an example, let us consider a constant vector field **A** in a plane and find its components in polar coordinates if we assume **A** to transform like a contravariant vector.

Let us measure θ from the direction of **A** as shown in Figure 12.1.

At a point P with coordinates (r, θ), the vectorial components of **A** in polar coordinates are

$(A\cos\theta, -A\sin\theta)$.

Making use of (12.11), we easily see that the **contravariant form** of the constant vector in polar coordinates will be

$$A^r = A\cos\theta, \;\; A^\theta = -A\frac{\sin\theta}{r}, \tag{12.19}$$

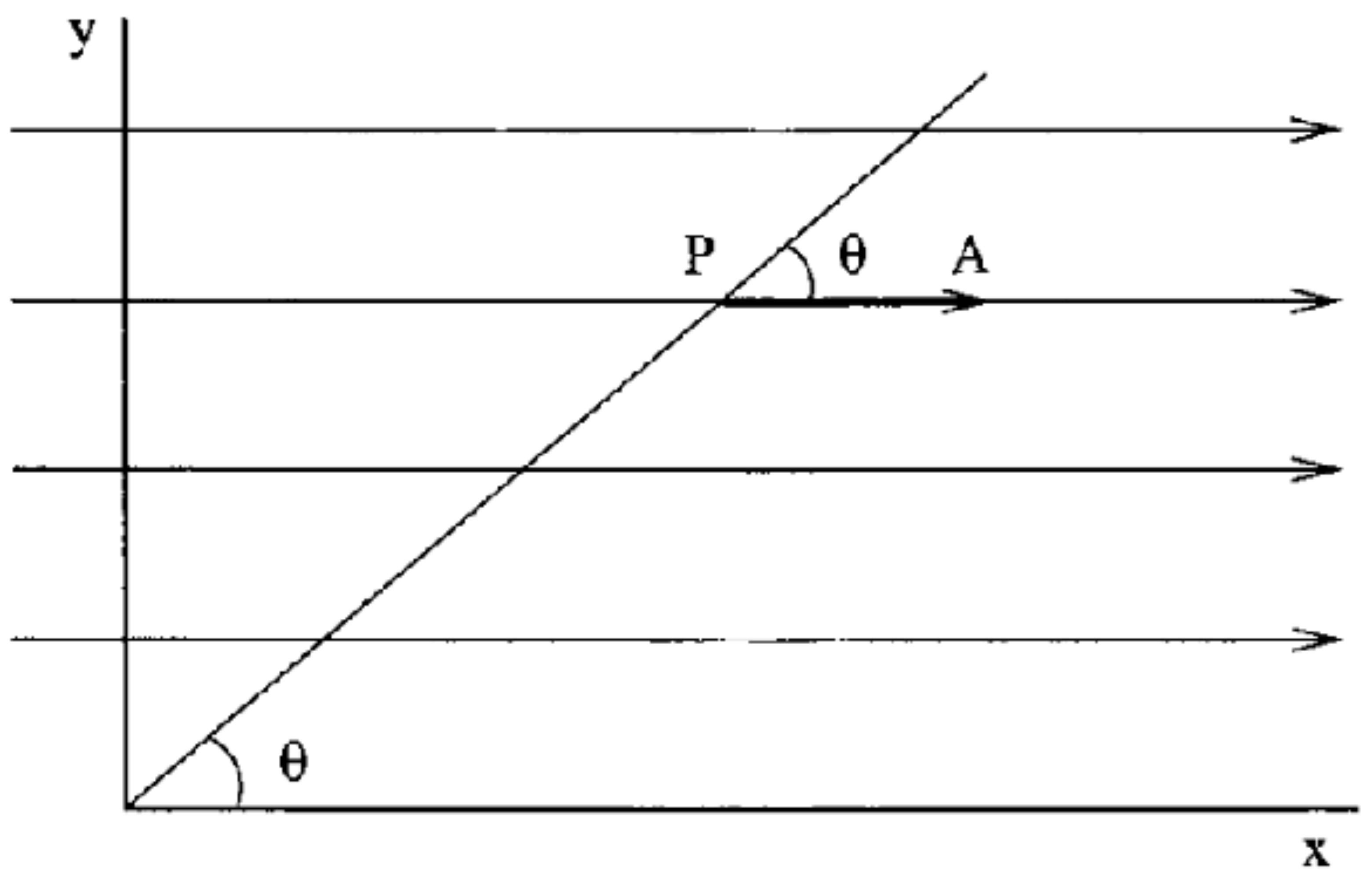

Fig. 12.1 A constant vector field **A**, shown with the point P where we consider the components of this vector field.

The metric tensor

whereas the **covariant form** will be

$$A_r = A\cos\theta, \;\; A_\theta = -Ar\sin\theta. \tag{12.20}$$

One can easily check that these expressions for different components will follow on applying (12.3) and (12.4) to the components in Cartesian coordinates: $A^x = A_x = A$, $A^y = A_y = 0$.

It should be apparent that **the question of whether a vector is contravariant or covariant does not make sense until we are told how its components in an arbitrary coordinate system are to be taken.**

For example, from a knowledge of how the generalized velocity or the generalized force is to be defined in different coordinates, we could use (12.1) and (12.2) to conclude that **they transform as contravariant and covariant vectors respectively.**

Differentiation of tensors

Let us consider a **contravariant vector field** $A^k(x^m)$. In **another coordinate system** $\bar{x}^l$, this same vector field will be denoted by $\bar{A}^i(\bar{x}^l)$, the transformation law being given by (12.3).

It readily follows from (12.3) that

$$\frac{\partial \overline{A}^i}{\partial \overline{x}^l} = \frac{\partial A^k}{\partial x^m} \frac{\partial x^m}{\partial \overline{x}^l} \frac{\partial \overline{x}^i}{\partial x^k} + A^k \frac{\partial x^m}{\partial \overline{x}^l} \frac{\partial^2 \overline{x}^i}{\partial x^m \partial x^k}.$$

Due to the presence of the **second term on the right-hand side**, it is clear that the derivative $\partial A^k / \partial x^m$ **does not transform like a tensor.**

To understand the physical reason behind it, we consider the **contravariant form of a constant vector field** given by (12.19). It follows from (12.19) that

$$\frac{\partial A^r}{\partial \theta} = -A \sin \theta.$$

Even though we may expect the derivative of a constant vector field to be zero, that is not the case.

Differentiation of tensors

This is presumably due to the fact that we are **using curvilinear coordinates**.

From the derivative, we have to remove the part coming due to the curvature of the coordinate system in order to get a more physically meaningful expression of the derivative. We now discuss how to do it.

In order to differentiate a vector field **A**, we need to subtract **A**(**x**) from **A**(**x** + d**x**).

Now we can sensibly talk of **adding or subtracting vectors only if they are at the same point**. So we need to do what is called a ***parallel transport*** of **A**(**x**) to **x** + d**x** before we subtract it from **A**(**x** + d**x**).

We expect that **A** will change to **A** + δ**A** under such a parallel transport.

For example, when we transport a vector **A** even on a plane surface, its components in polar coordinates (A^r , A^θ) will in general change. On physical grounds, **we expect that the change** δA^i in A^i **under parallel transport** from x^l to $x^l + dx^l$ **will be proportional to the displacement and to the vector itself.**

Hence we can write

$$\delta A^i = -\Gamma^i_{kl} A^k dx^l, \qquad (12.21)$$

where Γ^i_{kl} is known as the ***Christoffel symbol***. In mathematics and physics, the **Christoffel symbols are an array of numbers describing a metric connection.**

$$\delta A^i = -\Gamma^i_{kl} A^k dx^l, \tag{12.21}$$

Differentiation of tensors

The Christoffel symbol is not a tensor because δA^i is not a tensor.

We now expect a proper derivative of A^i to be given by

$$\frac{DA^i}{Dx^l} = \lim_{dx^l \to 0} \frac{A^i(x^l + dx^l) - [A^i(x^l) + \delta A^i]}{dx^l}.$$

On substituting for δA^i from (12.21) and writing

$$A^i(x^l + dx^l) = A^i(x^l) + \frac{\partial A^i}{\partial x^l} dx^l,$$

we get

$$\boxed{\frac{DA^i}{Dx^l} = \frac{\partial A^i}{\partial x^l} + \Gamma^i_{kl} A^k. \qquad (12.22)}$$

This is known as the ***covariant derivative*.**

Differentiation of tensors

It may be noted that the **ordinary derivative** $\partial A^i/\partial x^l$ and the **covariant derivative** DA^i/Dx^l **are sometimes denoted by symbols** $A^i_{,l}$ and $A^i_{;l}$. Other popular symbols for these derivatives are $\partial_l A^i$ and $\nabla_l A^i$.

We shall, use the longer notation here in order to avoid introducing too many new notations.

In order **to calculate the covariant derivate of a vector**, we first need to figure out **how to evaluate the Christoffel symbols** appearing in the expression (12.22) of the covariant derivative.

It is possible to find the Christoffel symbols from the metric tensor. For deriving the relation of Christoffel symbols to the metric tensor, we first have to figure out the expressions of covariant derivatives for a covariant vector A_i and tensors of higher rank.

By noting that $A_i B^i$ is a scalar for which we must have $\delta(A_i B^i) = 0$, we obtain

$$B^i \, \delta A_i = -A_i \, \delta B^i = A_i \, \Gamma^i_{kl} B^k \, dx^l .$$

$$B^i\,\delta A_i = -A_i\,\delta B^i = A_i\,\Gamma^i_{kl} B^k\,dx^l.$$

Differentiation of tensors

For indices which are summed (i.e. which are repeated twice above and below), we can change the symbols without affecting anything else. So we can write the above relation as

$$B^i\,\delta A_i = A_k \Gamma^k_{il}\,B^i\,dx^l,$$

from which it follows that

$$\delta A_i = \Gamma^k_{il}\,A_k\,dx^l. \tag{12.23}$$

It is now easy to show that

$$\frac{DA_i}{Dx^l} = \frac{\partial A_i}{\partial x^l} - \Gamma^k_{il}\,A_k. \tag{12.24}$$

The covariant derivative of the tensor A_{ik} must be given by

$$\frac{DA_{ik}}{Dx^l} = \frac{\partial A_{ik}}{\partial x^l} - \Gamma^m_{kl}\,A_{im} - \Gamma^m_{il}\,A_{mk}. \tag{12.25}$$

Differentiation of tensors

We now show that the Christoffel symbol is symmetric in its bottom two indices. Let us consider a vector

$$A_i = \frac{\partial V}{\partial x^i},$$

where V is a scalar field. We readily find that

$$\frac{DA_i}{Dx^k} - \frac{DA_k}{Dx^i} = (\Gamma^l_{ki} - \Gamma^l_{ik})\frac{\partial V}{\partial x^l}.$$

We note that the **left-hand side is a tensor,** which should transform between frames by obeying tensor transformation formulae. It is obvious that **the left-hand side in a Cartesian frame is zero and hence it must be zero in all frames**. Then the **right-hand side also should be zero**, implying

$$\boxed{\Gamma^l_{ik} = \Gamma^l_{ki}. \qquad (12.26)}$$

Differentiation of tensors

Now, if A_i is the covariant vector associated with A^i , then

$$\frac{DA_i}{Dx^l} = \frac{D}{Dx^l}(g_{ik}A^k) = g_{ik}\frac{DA^k}{Dx^l} + A^k\frac{Dg_{ik}}{Dx^l}.$$

Since A_i **and** A^i **essentially correspond to the same physical** entity, **their covariant derivatives also must be the same physical entity** and should be related to each other as

$$\frac{DA_i}{Dx^l} = g_{ik}\frac{DA^k}{Dx^l}.$$

It then follows that we must have

$$\frac{Dg_{ik}}{Dx^l} = 0. \tag{12.27}$$

Differentiation of tensors

From (12.25) and (12.27), we have

$$\frac{\partial g_{ik}}{\partial x^l} = \Gamma^m_{kl}\, g_{im} + \Gamma^m_{il}\, g_{mk}. \tag{12.28}$$

On permuting the symbols, we also have

$$\frac{\partial g_{li}}{\partial x^k} = \Gamma^m_{ik}\, g_{lm} + \Gamma^m_{lk}\, g_{mi}, \tag{12.29}$$

$$\frac{\partial g_{kl}}{\partial x^i} = \Gamma^m_{li}\, g_{km} + \Gamma^m_{ki}\, g_{ml}. \tag{12.30}$$

On subtracting (12.28) from the sum of (12.29) and (12.30), we get

$$2\Gamma^n_{ik} g_{ln} = \frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{kl}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l}$$

Differentiation of tensors

on keeping in mind the **symmetry property** (12.26) of the Christoffel symbol (as well as the symmetry property of the metric tensor). On multiplying this equation by g^{ml} and making use of (12.16), we finally have

$$\boxed{\Gamma^m_{ik} = \frac{1}{2} g^{ml} \left(\frac{\partial g_{li}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l} \right). \qquad (12.31)}$$

This is the **final expression of the Christoffel symbol in terms of the metric tensor.**

If we have the metric tensor for a space, we can **calculate the Christoffel symbols** by using (12.31) and then **work out any covariant derivatives.**

It is obvious that the **Christoffel symbols are zero in a Cartesian coordinate system**. Even when the **space is curved**, it can be shown that it is possible **to introduce Cartesian coordinates in a local region in such a way that the spatial derivatives of the metric tensor are zero and the Christoffel symbols vanish**. It is, however, not possible to make the higher derivatives of the metric tensor zero in a general situation.

Differentiation of tensors

$$ds^2 = dr^2 + r^2\, d\theta^2, \tag{12.8}$$

$$ds^2 = a^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{12.9}$$

$$g_{rr} = 1,\ g_{\theta\theta} = r^2,\ g_{r\theta} = g_{\theta r} = 0, \tag{12.11}$$

$$g_{\theta\theta} = a^2,\ g_{\phi\phi} = a^2 \sin^2\theta,\ g_{\theta\phi} = g_{\phi\theta} = 0, \tag{12.12}$$

Calculating Christoffel symbols for spaces of several dimensions can involve quite a bit of algebra, even though the algebra is usually straightforward if the metric tensor is not too complicated.

For four-dimensional spacetime, for example, ik in Γ^m_{ik} **can have 10 independent combinations** in view of the ik symmetry. Since $\boldsymbol{m}$ **can have four possible values**, it is clear that the **Christoffel symbol will have 40 components in four-dimensional spacetime**.

One has to do **further computations to obtain the curvature of space-time from the Christoffel symbols.** In general relativistic applications, often we have to do this very long algebra involving quantities with many components.

We shall only consider the two-dimensional metrics (12.8) and (12.9) here for illustrative purposes. The metric tensors for them are explicitly written down in (12.11) and (12.12) respectively. On substituting (12.11) into (12.31), we find that **the Christoffel symbols for the plane in polar coordinates** are

$$\Gamma^r_{\theta\theta} = -r,\ \Gamma^\theta_{r\theta} = \frac{1}{r}, \tag{12.32}$$

whereas the other four components (since the Christoffel symbol has six independent components in two dimensions) turn out to be zero.

Differentiation of tensors

For the surface of the sphere with the metric tensor given by (12.12), the only non-zero components are

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta, \;\; \Gamma^{\phi}_{\theta\phi} = \cot\theta. \qquad (12.33)$$

We have seen that the ordinary derivatives of the constant vector field given by (12.19) are non-zero.

Now we are ready to show that the **covariant derivatives are zero**, as we expect in the case of a constant vector field. Using (12.22), we can write down

$$\frac{DA^r}{D\theta} = \frac{\partial A^r}{\partial\theta} + \Gamma^r_{k\theta} A^k = \frac{\partial A^r}{\partial\theta} + \Gamma^r_{\theta\theta} A^\theta,$$

since the other term is zero due to the fact that $\Gamma^r_{t\theta} = 0$. On substituting for (A^r, A^θ) from (12.19) and using $\Gamma^r_{\theta\theta} = -r$, we find that

$$\frac{DA^r}{D\theta} = 0.$$

Differentiation of tensors

The other components of the covariant derivative can also similarly be shown to be zero.

Since the covariant derivative is a proper tensor and we know that the covariant derivative of a constant vector field in Cartesian coordinates (where the derivative reduces to an ordinary derivative) **is zero, we would expect it to be zero in the other coordinates as well.**

Curvature

Suppose a **vector A^i lying** on a flat surface is made to undergo **parallel transport along a closed path** and eventually brought back to its original location. We expect the final vector to be identical with the initial vector.

On the other hand, **if we have a curved surface, then the parallel transport of a vector along some arbitrary closed path may not bring it back to its original self.** From (12.23) we conclude that the change in the A_k after parallel transport along a closed path C will be

$$\Delta A_k = \oint_C \Gamma^i_{kl}\, A_i\, dx^l. \qquad (12.34)$$

Whether a **surface is plane or curved** can be inferred by finding out if the **right-hand side** of (12.34) is **zero or non-zero for arbitrary closed paths.** The same considerations should apply to higher dimensions as well.

We can conclude **a space to have zero curvature if** $\oint_C \Gamma^i_{kl} A_i dx^l$ around **any arbitrary closed path is always zero**. On the other hand, if we can find some closed path such that this line integral over it is **non-zero, then the space must be curved.**

$$\Delta A_k = \oint_C \Gamma^i_{kl} A_i \, dx^l. \tag{12.34}$$

Curvature

To proceed further, we have to **convert the line integral of (12.34) into a surface integral.**

Let us try to write down **Stokes's theorem of ordinary vector analysis** in tensorial notation. We consider ordinary **three-dimensional Cartesian space.** An element of area $d\mathbf{s}$ is a pseudovector from which we can construct a tensor

$$df^{ik} = \begin{pmatrix} 0 & ds_z & -ds_y \\ -ds_z & 0 & ds_x \\ ds_y & -ds_x & 0 \end{pmatrix}.$$

On using this notation, Stokes's theorem can be written as

$$\oint_C A_i \, dx^i = \frac{1}{2} \int df^{ik} \left(\frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k} \right), \tag{12.35}$$

Stokes's theorem: The line integral of a vector field over a loop is equal to the surface integral of its curl over the enclosed surface.

Curvature

$$\oint_C A_i\,dx^i = \frac{1}{2}\int df^{ik}\left(\frac{\partial A_k}{\partial x^i} - \frac{\partial A_i}{\partial x^k}\right), \tag{12.35}$$

where **the right-hand side is a surface integral** over a surface of which C is the boundary.

This tensorial expression of Stokes's theorem **holds for curved space and for higher dimensions also**.

On converting the right-hand side of (12.34) into a surface integral with the help of Stokes's theorem, we get

$$\Delta A_k = \frac{1}{2}\int\left[\frac{\partial}{\partial x^l}(\Gamma^i_{km}A_i) - \frac{\partial}{\partial x^m}(\Gamma^i_{kl}A_i)\right]df^{lm}$$

$$= \frac{1}{2}\int\left[\frac{\partial \Gamma^i_{km}}{\partial x^l}A_i - \frac{\partial \Gamma^i_{kl}}{\partial x^m}A_i + \Gamma^i_{km}\frac{\partial A_i}{\partial x^l} - \Gamma^i_{kl}\frac{\partial A_i}{\partial x^m}\right]df^{lm}.$$

Now the change in A_i in the present situation is caused by parallel transport from its original position. So the change in A_i must be given by (12.23) from which it follows that

$$\frac{\partial A_i}{\partial x^l} = \Gamma^n_{il}\,A_n.$$

Curvature

Using this we get

$$\Delta A_k = \frac{1}{2}\int \left[\frac{\partial \Gamma^i_{km}}{\partial x^l} - \frac{\partial \Gamma^i_{kl}}{\partial x^m} + \Gamma^n_{km}\Gamma^i_{nl} - \Gamma^n_{kl}\Gamma^i_{nm}\right] A_i\, df^{lm}.$$

This can be written as

$$\Delta A_k = \frac{1}{2}\int R^i_{klm}\, A_i\, df^{lm}, \tag{12.36}$$

where

$$R^i_{klm} = \frac{\partial \Gamma^i_{km}}{\partial x^l} - \frac{\partial \Gamma^i_{kl}}{\partial x^m} + \Gamma^n_{km}\Gamma^i_{nl} - \Gamma^n_{kl}\Gamma^i_{nm}. \tag{12.37}$$

Curvature

As we have already pointed out, whether a space is flat or **curved can be determined by finding out if ΔA_k is always zero or not**. This in turn depends of **whether R^i_{kml} is zero or non-zero**, as should be clear from (12.36).

We thus conclude that R^i_{kml} given by (12.37) **is a measure of the curvature of space**. It is called the ***Riemann curvature tensor.***

Just from the definition of R^i_{kml} given by (12.37), the following **symmetry properties follow**

$$R^i_{klm} = -R^i_{kml}, \tag{12.38}$$

$$R^i_{klm} + R^i_{mkl} + R^i_{lmk} = 0. \tag{12.39}$$

Another important result, known as the ***Bianchi identity***, is

$$\frac{DR^n_{ikl}}{Dx^m} + \frac{DR^n_{imk}}{Dx^l} + \frac{DR^n_{ilm}}{Dx^k} = 0. \tag{12.40}$$

Curvature

The best way of proving this identity is to go to a Cartesian frame and use the result following from (12.5) that, **if a tensor has all components zero in a Cartesian frame, it must be identically zero in all frames.**

Now, **if the space under consideration is curved, it is possible to introduce a Cartesian coordinate only in a local region**. The Christoffel symbol Γ^i_{km} can be made zero in the local Cartesian frame, but its derivatives may not be zero.

It follows from (12.37) that **at this local point** we should have

$$\frac{DR^n_{ikl}}{Dx^m} = \frac{\partial^2 \Gamma^n_{il}}{\partial x^m \partial x^k} - \frac{\partial^2 \Gamma^n_{ik}}{\partial x^m \partial x^l},$$

since the covariant derivative reduces to the ordinary derivative in the local region of the Cartesian frame. If we write similar expressions for the other terms in (12.40) and add them up, we establish the identity in the local Cartesian frame.

Due to its tensorial nature, it then follows that the Bianchi identity must be a general identity true in all frames.

Curvature

From the curvature tensor R^i_{kml} , we can obtain a tensor R_{km} of lower rank by contracting i with l:

$$\boxed{R_{km} = R^i_{kim} = \frac{\partial \Gamma^i_{km}}{\partial x^i} - \frac{\partial \Gamma^i_{ki}}{\partial x^m} + \Gamma^n_{km}\Gamma^i_{ni} - \Gamma^n_{ki}\Gamma^i_{nm}. \qquad (12.41)}$$

This tensor R_{km} **is known as the *Ricci tensor***. It is straightforward to show that R_{km} is **symmetric,** i.e.

$$R_{km} = R_{mk}.$$

We can finally obtain a scalar

$$\boxed{R = g^{mk} R_{mk} \qquad (12.42)}$$

known as the *scalar curvature*.

Curvature

Now we can apply this to the **metric (12.9) pertaining to the surface of a sphere,** of which the non-zero Christoffel symbols are given in (12.33).

Let us try to **calculate the scalar curvature,** for which we need $R_{\theta\theta}$ and $R_{\phi\phi}$. Now

$$R_{\theta\theta} = R^{\theta}_{\theta\theta\theta} + R^{\phi}_{\theta\phi\theta} = R^{\phi}_{\theta\phi\theta},$$

since $R^{\theta}_{\theta\theta\theta} = 0$ due to the antisymmetry property (12.38). We thus **need to calculate only one component of the curvature tensor to obtain $R_{\theta\theta}$.**

On using the expressions of Christoffel symbols given by (12.33), it easily follows from the definition (12.37) of the curvature tensor that

$$R_{\theta\theta} = R^{\phi}_{\theta\phi\theta} = 1.$$

A similar calculation gives

$$R_{\phi\phi} = R^{\theta}_{\phi\theta\phi} = \sin^2\theta.$$

$$\boxed{R = g^{mk} R_{mk} \qquad (12.42)}$$

Curvature

Finally **the scalar curvature is given by**

$$R = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = \frac{1}{a^2}.1 + \frac{1}{a^2 \sin^2\theta}.\sin^2\theta = \frac{2}{a^2} \qquad (12.43)$$

on making use of (12.17).

Although the Riemann tensor has many components, we did not have to do too much algebra to obtain the scalar curvature.

Thus calculating the curvature of a simple two-dimensional metric is not too complicated. The **curvature calculation for any four-dimensional spacetime metric usually involves a large amount of straightforward algebra**.

Curvature

We also note from (12.43) that **the curvature scalar is a constant over the surface of the sphere,** which is expected from the fact that this surface is uniform.

The scalar curvature for the metric (12.10) is $-2/a^2$.

The **metrics (12.9) and (12.10) give the only two uniformly curved surfaces possible in two dimensions**, one with uniform positive curvature and the other with uniform negative curvature.

It is easy to show that all components of the curvature tensor are zero for the metric (12.8) corresponding to a plane. **Given the metric of a space (or a spacetime), we now have the methods to find out whether it is curved or flat.**

We end our discussion of curvature by introducing another tensor which turns out to be very important in the formulation of general relativity, as we shall see later. It is the ***Einstein tensor*** defined as

$$\boxed{G_{ik} = R_{ik} - \frac{1}{2} g_{ik} R. \qquad (12.44)}$$

Curvature

One of its very important properties is that it is a **divergenceless tensor** satisfying

$$\frac{DG_{ik}}{Dx_k} = 0. \qquad (12.45)$$

From the definition (12.44) of the Einstein tensor, this implies

$$\frac{DR_{ik}}{Dx_k} - \frac{1}{2} g_{ik} \frac{\partial R}{\partial x_k} = 0. \qquad (12.46)$$

To prove this, we need to begin from the **Bianchi identity** (12.40).

Writing $R^n_{imk} = -R^n_{imk}$ in the second term of (12.40), we

(i) contract n and k, and

(ii) multiply by g^{im}. This gives

$$\frac{DR^m_l}{Dx^m} - \frac{\partial R}{\partial x^l} + \frac{D}{Dx^n}(g^{im} R^n_{ilm}) = 0 \qquad (12.47)$$

Curvature

on remembering that the covariant derivative of the metric tensor is zero, as indicated by (12.27), and the covariant derivative of a scalar reduces to its ordinary derivative.

$g^{im} R^n_{ilm} = R^n_l$ which implies that the last term in (12.47) is the same as the first term, giving an equation which is easily seen to be equivalent with (12.46). This essentially establishes that **the Einstein tensor is divergenceless** as encapsulated in (12.45) – a result which is going to be crucially **important when we formulate general relativity.**

Geodesics

The shortest path between two points in a plane surface or in a flat space is a straight line.

If the surface or the space is curved, then the shortest path between two points is called a *geodesic*.

One of the central ideas of **general relativity** is that **a particle moves along a geodesic in the four-dimensional spacetime**.

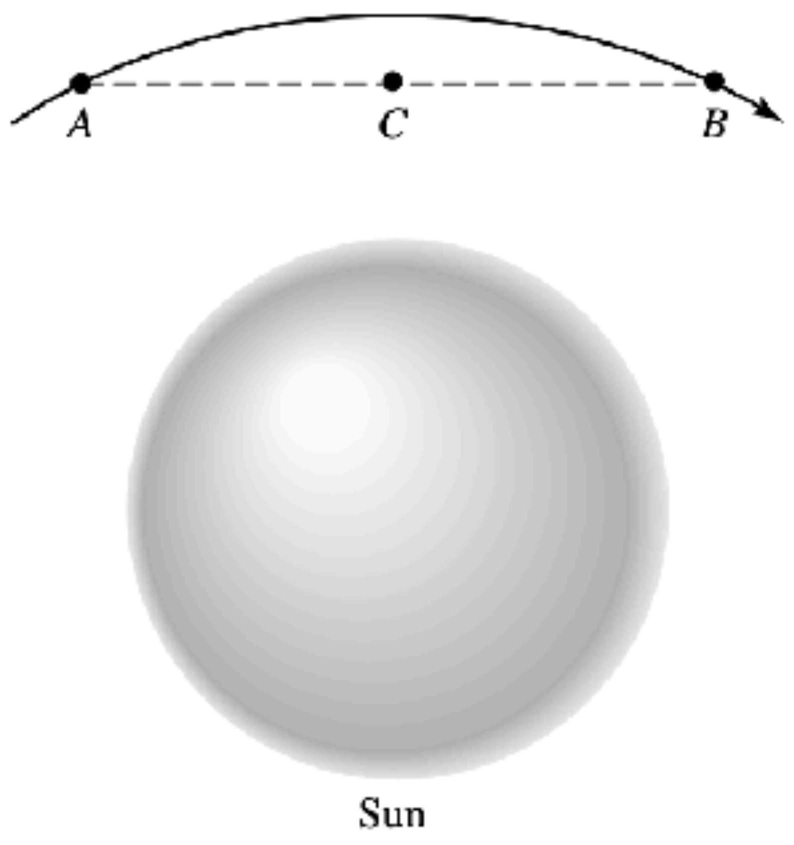

FIGURE 3 A photon's path around the Sun is shown by the solid line. The bend in the photon's trajectory is greatly exaggerated.

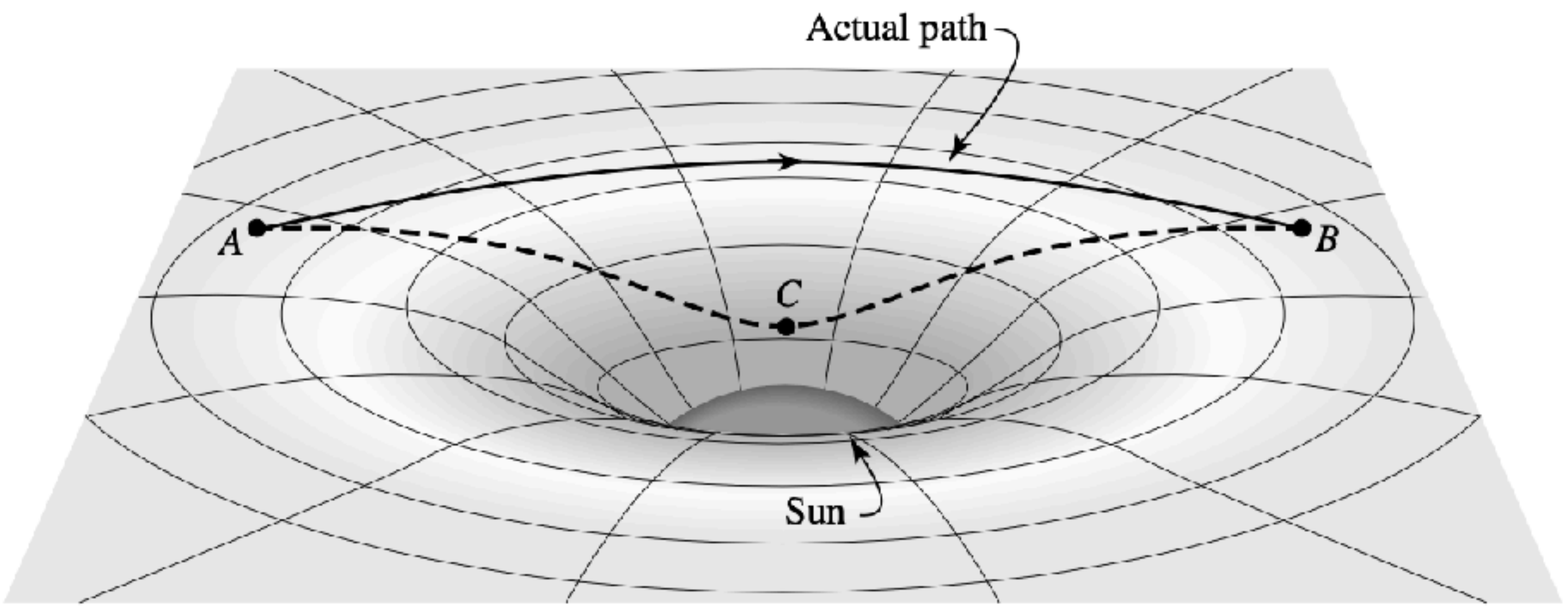

FIGURE 4 Comparison of two photon paths through curved space between points A and B. The projection of the path ACB onto the plane is the straight line depicted in Fig. 3.

Geodesics

The last mathematical question we have to address is to show **how we obtain geodesics in a particular space of which the metric tensor is known.**

Let us consider an **arbitrary path between two points A and B**. The **length ds** of a small segment of this path is given by (12.7). Hence the **length of the whole path** must be

$$s = \int_A^B \sqrt{g_{ik} \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}}\, d\lambda, \tag{12.48}$$

where λ is a parameter measured along the path.

Simplifying the inside of the integral

$$L = \sqrt{g_{ik} \frac{dx^i}{d\lambda} \frac{dx^k}{d\lambda}} \tag{12.49}$$

Geodesics

the **length of the path is**

$$s = \int_A^B L\, d\lambda$$

and **the condition for the path to be an extremum is given by the Lagrange equation** (we are looking for the shortest path)

$$\boxed{\frac{d}{d\lambda}\left(\frac{\partial L}{\partial(dx^i/d\lambda)}\right) - \frac{\partial L}{\partial x^i} = 0.}$$

On substituting the expression of L given by (12.49) into this Lagrange equation and remembering that

$$\sqrt{g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda}} = \frac{ds}{d\lambda},$$

Geodesics

a few steps of algebra give

$$\frac{d}{ds}\left(g_{ik}\frac{dx^k}{ds}\right) - \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\frac{dx^k}{ds}\frac{dx^l}{ds} = 0. \qquad (12.50)$$

Now the first term of this equation is

$$\begin{aligned}\frac{d}{ds}\left(g_{ik}\frac{dx^k}{ds}\right) &= g_{ik}\frac{d^2x^k}{ds^2} + \frac{\partial g_{ik}}{\partial x^l}\frac{dx^l}{ds}\frac{dx^k}{ds} \\ &= g_{ik}\frac{d^2x^k}{ds^2} + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^l} + \frac{\partial g_{il}}{\partial x^k}\right)\frac{dx^l}{ds}\frac{dx^k}{ds}.\end{aligned}$$

On **substituting** this in (12.50), we have

$$g_{ik}\frac{d^2x^k}{ds^2} = -\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^l} + \frac{\partial g_{il}}{\partial x^k} - \frac{\partial g_{kl}}{\partial x^i}\right)\frac{dx^k}{ds}\frac{dx^l}{ds}.$$

Geodesics

Multiplying this by g^{mi} , we finally get

$$\boxed{\frac{d^2x^m}{ds^2} = -\Gamma^m_{kl}\frac{dx^k}{ds}\frac{dx^l}{ds}, \qquad (12.51)}$$

where Γ^m_{kl} is the Christoffel symbol defined in (12.31).

This equation (12.51) **is the** ***geodesic equation*** which **has to be satisfied by a curve in space if it happens to be a geodesic in that space.**
Let's illustrate this by **applying it to one of the two-dimensional metrics.**

Since the metric (12.8) corresponds to a plane surface, the geodesic for this metric must be a straight line. Hence we expect that the geodesic equation (12.51) applied to this metric should give us a straight line. We now explicitly show this.

Geodesics

The non-zero Christoffel symbols of this metric (12.8) are given in (12.32). On substituting these in (12.51), we get the following two equations

$$\frac{d^2r}{ds^2} = r\left(\frac{d\theta}{ds}\right)^2, \tag{12.52}$$

$$\frac{d^2\theta}{ds^2} = -\frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds}. \tag{12.53}$$

The factor 2 in (12.53) comes from the two equal terms arising out of the combinations $r\theta$ and θr. (12.53) is equivalent to

$$\frac{1}{r^2}\frac{d}{ds}\left(r^2\frac{d\theta}{ds}\right) = 0.$$

Geodesics

It follows from this that $r^2(d\theta/ds)$ **must be a constant along the geodesic**. We therefore write

$$\frac{d\theta}{ds} = \frac{l}{r^2}, \tag{12.54}$$

where l is a constant. On **dividing the metric (12.8) by ds^2**, we obtain

$$\left(\frac{dr}{ds}\right)^2 = 1 - r^2\left(\frac{d\theta}{ds}\right)^2 = 1 - \frac{l^2}{r^2}$$

so that

$$\frac{dr}{ds} = \pm\sqrt{1 - \frac{l^2}{r^2}}. \tag{12.55}$$

Geodesics

Dividing (12.54) by (12.55), we get

$$\frac{d\theta}{dr} = \pm\frac{l/r^2}{\sqrt{1-\frac{l^2}{r^2}}},$$

of which a **solution** is

$$\theta = \theta_0 \pm \cos^{-1}\left(\frac{l}{r}\right)$$

where θ_0 **is the constant of integration**.

This solution can be put in the form

$$r\cos(\theta-\theta_0) = l, \tag{12.56}$$

which is clearly the equation of **a straight line.**

This completes our proof that the geodesic in a plane is a straight line, even though it is mathematically not so apparent when we use polar coordinates.