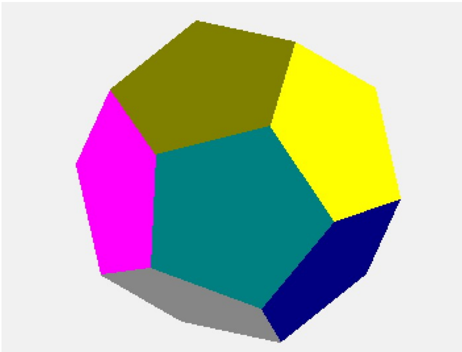
Differential Geometry
and
Tensor analysis
With applications to General Relativity

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



Ole Witt-Hansen 2015

Contents

1. Linear algebra

The concept of a tensor is much easier to grasp, if you have a solid background in linear algebra. So we start out with, some fundamental issues in that field.

A linear vector function is a vector function which obeys the following condition.

(1.1) $$f(\lambda\vec{a} + \mu\vec{b}) = \lambda f(\vec{a}) + \mu f(\vec{b}),$$

The relation being valid for arbitrary vectors $\vec{a}$ *and* $\vec{b}$ and arbitrary real numbers λ *and* μ.
If we have a base on the vector space consisting of mutually orthogonal unit vectors:

$$\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots.., \vec{e}_n$$

The linear function is completely determined by the mapping of these vectors, since a vector

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + \ldots. + x_n\vec{e}_n$$

As a consequence of (1.1) will be mapped into

(1.2) $$f(\vec{x}) = x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) + x_3 f(\vec{e}_3) + \ldots. + x_n f(\vec{e}_n)$$

Each of the vectors $f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3), \ldots., f(\vec{e}_n)$

can then be written as a linear combination of the basic vectors.

(1.3) $$f(\vec{e}_k) = a_{1k}\vec{e}_1 + a_{2k}\vec{e}_2 + a_{3k}\vec{e}_3 + \ldots..a_{nk}\vec{e}_n$$

I follows (because the base vectors are mutually orthogonal), that

$$a_{jk} = \vec{e}_j \cdot f(\vec{e}_k)$$

(1.3) is usually written in matrix form

(1.3) $$(f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3), \ldots., f(\vec{e}_n)) = (\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots.., \vec{e}_n)\begin{pmatrix} a_{11} & a_{12} & a_{13} & . & a_{1n} \\ a_{21} & a_{22} & a_{23} & . & a_{2n} \\ a_{31} & a_{32} & a_{33} & . & a_{3n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & a_{n3} & . & a_{nn} \end{pmatrix}$$

Or we can write the above relation in tensor form using Einstein's summation convention:

That whenever the same index appears twice, summation is implied.

(1.4) $$f(\vec{e}_k) = \sum_{j=1}^{n} a_{jk}\vec{e}_j$$ is written as

$$f(\vec{e}_k) = a_{jk}\vec{e}_j$$ (Summation *j = 1…n* is implied)

If $\vec{y} = f(\vec{x})$ we get from (1.2) and (1.3) using tensor notation:

$$f(\vec{x}) = x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) + x_3 f(\vec{e}_3) + ... + x_n f(\vec{e}_n) = x_j f(\vec{e}_j)$$

And

$$f(\vec{e}_k) = a_{1k}\vec{e}_1 + a_{2k}\vec{e}_2 + a_{3k}\vec{e}_3 + ...a_{nk}\vec{e}_n = a_{ik}\vec{e}_i$$

We get: $\vec{y} = f(\vec{x}) = a_{ij}x_j\vec{e}_i$

Which implies: $y_i = a_{ij}x_j$

Or when written in matrix form:

(1.5)
$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & . & a_{1n} \\ a_{21} & a_{22} & a_{23} & . & a_{2n} \\ a_{31} & a_{32} & a_{33} & . & a_{3n} \\ . & . & . & . & . \\ a_{n1} & a_{n2} & a_{n3} & . & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ x_n \end{pmatrix}$$

And when written in symbolic matrix form.

$$\underline{y} = \underline{\underline{A}}\ \underline{x}$$

Where we use double underscore to mark a matrix symbol.

The multiplication of two matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ with elements a_{ij} and b_{ij} to give the matrix $\underline{\underline{C}}$ with elements c_{ij} goes as follows: One makes the "scalar product" of the i'th row in $\underline{\underline{A}}$ and the *j*'th column in $\underline{\underline{B}}$. Written in tensor notation:

(1.6) $$c_{ij} = a_{ik}b_{kj}$$ (implied summation over *k*)

The determinant of a matrix $\underline{\underline{A}}$ is written *det*($\underline{\underline{A}}$) or |$\underline{\underline{A}}$|

The determinant of a 2x2 matrix:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

Whereas the determinant of a 3x3 matrix is:

(1.7) $$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{12}a_{33} - a_{21}a_{32}a_{23} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}$$

The rule is that each factor in the product $a_{1i}a_{2j}a_{3k}$ belongs to a different row and a different column from the other factors. If [*i, j, k*] is an even permutation of [1*, 2, 3*] then the term is signed with a plus, if it is an uneven permutation, the term is signed with a minus.

The determinant, can also be expressed with the help of the Levi-Civitas symbol

$$\varepsilon_{ijk} = \begin{vmatrix} 1 & \textit{if } \; i,j,k \; \textit{is an even permutation of } \; 1,2,3 \\ -1 & \textit{if } \; i,j,k \; \textit{is an uneven permutation of } \; 1,2,3 \\ 0 & \textit{otherwise} \end{vmatrix}$$

The determinant of the matrix shown above can then be written:

$$det(\underline{\underline{A}}) = a_{1i}a_{2j}a_{3k}\varepsilon_{ijk}$$

Since there are six permutations of three elements, the number of terms is six.

The extension to higher dimensions including the Levi-Civitas symbol is straightforward.

The unit matrix $\underline{\underline{E}}$ has the elements δ_{ij}, where δ_{ij} is the Koneke symbol

(1.7) $$\delta_{ij} = 1 \text{ if } i = j \text{ else } 0$$

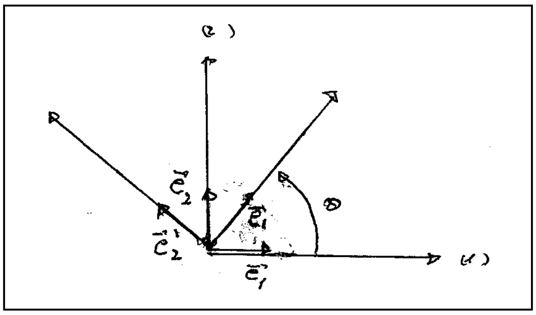
Any *n x n* matrix that has a non zero determinant, is called regular. Any regular matrix $\underline{\underline{A}}$ has an inverse matrix $\underline{\underline{A}}^{-1}$, defined by the equations:

(1.8) $$\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{E}}$$

If the elements in the inverse matrix are b_{ij}, then (1.7) can be written

$$a_{ik}b_{kj} = \delta_{ij}$$

1.1 Coordinate transformations



First we shall consider an ordinary coordinate system in the plane, which is rotated an angle θ.
The base vectors $(\vec{e}_1, \vec{e}_2)$ are rotated by an angle θ into $(\vec{e}_1', \vec{e}_2')$. From the figure we see, that:
$\vec{e}_1' = \vec{e}_1 \cos\theta + \vec{e}_2 \sin\theta$ *and* $\vec{e}_2' = -\vec{e}_1 \sin\theta + \vec{e}_2 \cos\theta$

When written in matrix form.

$$(\vec{e}_1{}', \vec{e}_2{}') = (\vec{e}_1, \vec{e}_2)\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

According to (1.3) and (1.5), we get the corresponding transformation for the coordinates (x_1 , x_2) and (x_1' , x_2').

(1.9)
$$\begin{pmatrix} x_1{}' \\ x_2{}' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

In 3-dimensional space, this transformation corresponds to a rotation around the z-axis, an angle θ. The transformation matrix then becomes:

(1.10)
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
(Rotation about the z-axis)

For the rotations around the y-axis or the x-axis, we have quite similar expressions

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
(Rotation about the y-axis)

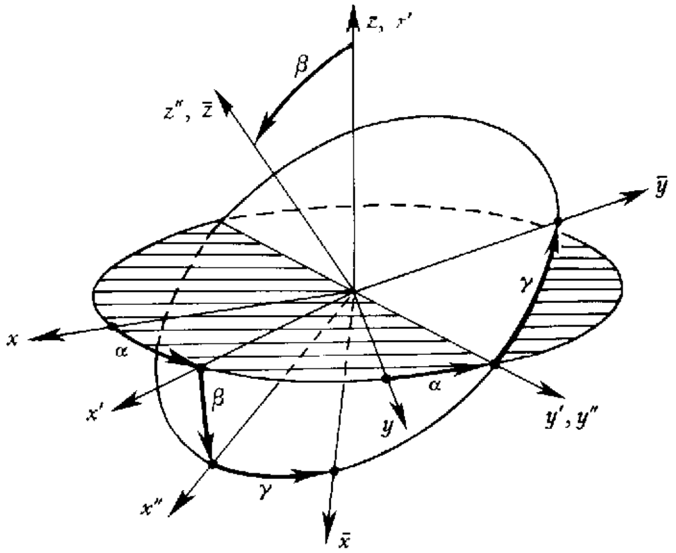
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
(Rotation about the x-axis)

To bring the coordinate system into an arbitrary angular position, we need three rotations.
They are traditionally chosen as a rotation around the *z*-axis an angle α, followed by a rotation around the new *y*-axis (y') an angle β, and finally an angle γ around the new *z*-axis z''.
The overall rotation is found by multiplying the three matrices.
The angles α, β, γ are called the Euler angles. They are illustrated below:

$$R(\alpha,\beta,\gamma) = R_z{}''(\gamma)\, R_y{}'(\beta)\, R_z(\alpha) =$$
$$\begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$\begin{pmatrix} \cos\beta\cos\alpha\cos\gamma - \sin\alpha\sin\gamma & \cos\beta\sin\alpha\cos\gamma + \cos\alpha\sin\gamma & -\sin\beta\cos\gamma \\ -\cos\beta\cos\alpha\sin\gamma - \sin\alpha\cos\gamma & -\cos\beta\sin\alpha\sin\gamma + \cos\alpha\cos\gamma & \sin\beta\sin\gamma \\ \sin\beta\cos\alpha & \sin\beta\sin\alpha & \cos\beta \end{pmatrix}$$

Showing the Euler angles.



2. Generalized coordinates

Hitherto we have considered only Cartesian coordinates, with an orthonormal base $\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots\ldots, \vec{e}_n$. This means that the base fulfils the condition $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$. The basic vectors are the same for every point in space. Because of the orthogonality the square of the infinitesimal distance element is the same in every point.

$$d\vec{s} = \vec{e}_1 dx_1 + \vec{e}_2 dx_2 + \vec{e}_3 dx_3 + \ldots. + \vec{e}_n dx_n$$

(2.1) $$ds^2 = d\vec{s} \cdot d\vec{s} = dx_1^2 + dx_2^2 + dx_3^2 + \ldots dx_n^2$$

In a generalized coordinate system $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n)$, the base vectors are not necessarily an orthogonal system, and the base may (and usually does) wary from point to point.
If $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n)$ represents a point in the generalized coordinates, we can define a base: $(\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots \vec{e}_n)$ along the axis of the coordinate system $\bar{x}$. A differential displacement is given by:

$$d\bar{x} = (d\bar{x}_1, d\bar{x}_2, d\bar{x}_3, \ldots, d\bar{x}_n)$$

And the displacement vector

(2.2) $$d\vec{s} = d\bar{x}_1 \vec{e}_1 + d\bar{x}_2 \vec{e}_2 + d\bar{x}_3 \vec{e}_3 + \ldots + d\bar{x}_n \vec{e}_n$$

The length of the distance element is:

$$ds^2 = d\vec{s}^{\,2} = d\vec{s} \cdot d\vec{s} = (d\bar{x}_1 \vec{e}_1 + d\bar{x}_2 \vec{e}_2 + d\bar{x}_3 \vec{e}_3 + \ldots + d\bar{x}_n \vec{e}_n) \cdot (d\bar{x}_1 \vec{e}_1 + d\bar{x}_2 \vec{e}_2 + d\bar{x}_3 \vec{e}_3 + \ldots + d\bar{x}_n \vec{e}_n)$$

(2.3) $$ds^2 = \vec{e}_i \cdot \vec{e}_j d\bar{x}_i d\bar{x}_j$$ (summation over *i* and *j* is understood)

There is a tradition, which we shall substantiate presently, letting the index on the coordinate functions go upstairs. From now on we write (2.3) as:

(2.3) $$ds^2 = \vec{e}_i \cdot \vec{e}_j d\bar{x}^i d\bar{x}^j$$

Since the base vectors in the generalized coordinate system, are not (necessarily) orthogonal vectors, the scalar products $\vec{e}_i \cdot \vec{e}_j$ does not vanish in general for $i \neq j$.
We define the metric (fundamental) form $\bar{g}_{ij} = \bar{g}_{ij}(\bar{x}) = \bar{g}_{ij}(\bar{x}_1, \bar{x}_2, \bar{x}_3, ..., \bar{x}_n)$ as

(2.4) $$\bar{g}_{ij} = \vec{e}_i \cdot \vec{e}_j$$

And the distance element becomes

$$d\bar{s}^2 = \bar{g}_{ij} d\bar{x}^i d\bar{x}^j$$

In the coordinate system $x = (x_1, x_2, x_3, ..., x_n)$ we write of course:

(2.4) $$g_{ij} = \vec{e}_i \cdot \vec{e}_j \quad \text{and} \quad ds^2 = g_{ij} dx^i dx^j$$

In a Cartesian coordinate system $g_{ij} = \vec{e}_i \cdot \vec{e}_j = \delta_{ij}$

It is important to note that although the relation between generalized coordinate are (in general) not linear relations, the relations between the differentials $d\bar{x} = (d\bar{x}_1, d\bar{x}_2, d\bar{x}_3, ..., d\bar{x}_n)$ are in fact linear. If

$$y = f(x_1, x_2, x_3, ..., x_n)$$

Then

(2.6) $$dy = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 + ... + \frac{\partial f}{\partial x_n} dx_n$$

Is in fact a linear form on the variables dx_i

The transformation to the coordinates $\bar{x}_i = \bar{x}_i(x_1, x_2, x_3, ..., x_n)$ goes as follows:

(2.7) $$d\bar{x}_k = \frac{\partial \bar{x}_k}{\partial x_1} dx_1 + \frac{\partial \bar{x}_k}{\partial x_2} dx_2 + \frac{\partial \bar{x}_k}{\partial x_3} dx_3 + ... + \frac{\partial \bar{x}_k}{\partial x_n} dx_n$$

Or in tensor notation with the indices raised:

(2.8) $$d\bar{x}^k = \frac{\partial \bar{x}^k}{\partial x^i} dx^i$$ (summation over index *i* is implied)

Similarly we may write the inverse transformation $x_i = x_i(\bar{x}_1, \bar{x}_2, \bar{x}_3, ..., \bar{x}_n)$

(2.9) $$dx^j = \frac{\partial x^j}{\partial \bar{x}^k} d\bar{x}^k$$

The transformation between the two coordinate differentials can also be written in matrix form:

(2.10) $$\begin{pmatrix} d\bar{x}_1 \\ d\bar{x}_2 \\ d\bar{x}_3 \\ \\ d\bar{x}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \frac{\partial \bar{x}_1}{\partial x_2} & \frac{\partial \bar{x}_1}{\partial x_3} & . & \frac{\partial \bar{x}_1}{\partial x_n} \\ \frac{\partial \bar{x}_2}{\partial x_1} & \frac{\partial \bar{x}_2}{\partial x_2} & \frac{\partial \bar{x}_2}{\partial x_3} & . & \frac{\partial \bar{x}_2}{\partial x_n} \\ \frac{\partial \bar{x}_3}{\partial x_1} & \frac{\partial \bar{x}_3}{\partial x_2} & \frac{\partial \bar{x}_3}{\partial x_3} & . & \frac{\partial \bar{x}_3}{\partial x_n} \\ . & . & . & . & . \\ \frac{\partial \bar{x}_n}{\partial x_1} & \frac{\partial \bar{x}_n}{\partial x_2} & \frac{\partial \bar{x}_n}{\partial x_3} & . & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \\ \\ dx_n \end{pmatrix}$$

The determinant of the transformation matrix $\underline{\underline{T}}$ $det(\underline{\underline{T}})$ is called the *Jacobian*, and it is the scale factor between the two volume elements: $d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 \cdots d\bar{x}_n$ and $dx_1 dx_2 dx_3 \cdots dx_n$.
This can be seen if we write the two vector columns as matrices.

(2.10) $$\begin{pmatrix} d\bar{x}_1 & 0 & 0 & 0 & 0 \\ 0 & d\bar{x}_2 & 0 & 0 & 0 \\ 0 & 0 & d\bar{x}_3 & 0 & 0 \\ 0 & 0 & 0 & . & 0 \\ 0 & 0 & 0 & 0 & d\bar{x}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \frac{\partial \bar{x}_1}{\partial x_2} & \frac{\partial \bar{x}_1}{\partial x_3} & . & \frac{\partial \bar{x}_1}{\partial x_n} \\ \frac{\partial \bar{x}_2}{\partial x_1} & \frac{\partial \bar{x}_2}{\partial x_2} & \frac{\partial \bar{x}_2}{\partial x_3} & . & \frac{\partial \bar{x}_2}{\partial x_n} \\ \frac{\partial \bar{x}_3}{\partial x_1} & \frac{\partial \bar{x}_3}{\partial x_2} & \frac{\partial \bar{x}_3}{\partial x_3} & . & \frac{\partial \bar{x}_3}{\partial x_n} \\ . & . & . & . & . \\ \frac{\partial \bar{x}_n}{\partial x_1} & \frac{\partial \bar{x}_n}{\partial x_2} & \frac{\partial \bar{x}_n}{\partial x_3} & . & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} dx_1 & 0 & 0 & 0 & 0 \\ 0 & dx_2 & 0 & 0 & 0 \\ 0 & 0 & dx_3 & 0 & 0 \\ 0 & 0 & 0 & . & 0 \\ 0 & 0 & 0 & 0 & dx_n \end{pmatrix}$$

It is easy to convince one, that this represent the same relation between the differentials as before. From algebra we know that the determinant of a diagonal matrix is the product of the diagonal elements. Also we know that the determinant of the matrix products of two matrices is the product of the two determinants:

$$det(\underline{\underline{A}}\ \underline{\underline{B}}) = det(\underline{\underline{A}})det(\underline{\underline{B}})$$

The assertion then follows, if we take the determinant of both sides.

(2.11) $$d\bar{x}_1 d\bar{x}_2 d\bar{x}_3 \cdots d\bar{x}_n = \det\left\{\frac{\partial \bar{x}_i}{\partial x_j}\right\} dx_1 dx_2 dx_3 \cdots dx_n$$

If $(x_1, x_2, x_3, ..., x_n)$ are Cartesian coordinates we have according to (2.1)

$$ds^2 = d\bar{s} \cdot d\bar{s} = dx^1 dx^1 + dx^2 dx^2 + dx^3 dx^3 + ... dx^n dx^n = dx^i dx^i$$

Inserting (2.9) $dx^j = \frac{\partial x^j}{\partial \bar{x}^k} d\bar{x}^k$ gives

$$ds^2 = (\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\partial x^i}{\partial \bar{x}^k} d\bar{x}^k)(\sum_{j=1}^{n}\sum_{l=1}^{n} \frac{\partial x^j}{\partial \bar{x}^l} d\bar{x}^l)$$

$$ds^2 = \sum_{k=1}^{n}\sum_{l=1}^{n}(\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial x^i}{\partial \bar{x}^k}\frac{\partial x^j}{\partial \bar{x}^l}) d\bar{x}^k d\bar{x}^l$$

From which it is seen that the metric form for the transformed coordinates is:

(2.10) $$\bar{g}_{kl} = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial x^i}{\partial \bar{x}^k}\frac{\partial x^j}{\partial \bar{x}^l}$$

The transformation formulas (2.8) and (2.9) can be written in matrix form:

(2.11) $$\begin{pmatrix} d\bar{x}_1 \\ d\bar{x}_2 \\ d\bar{x}_3 \\ . \\ d\bar{x}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \frac{\partial \bar{x}_1}{\partial x_2} & \frac{\partial \bar{x}_1}{\partial x_3} & . & \frac{\partial \bar{x}_1}{\partial x_n} \\ \frac{\partial \bar{x}_2}{\partial x_1} & \frac{\partial \bar{x}_2}{\partial x_2} & \frac{\partial \bar{x}_2}{\partial x_3} & . & \frac{\partial \bar{x}_2}{\partial x_n} \\ \frac{\partial \bar{x}_3}{\partial x_1} & \frac{\partial \bar{x}_3}{\partial x_2} & \frac{\partial \bar{x}_3}{\partial x_3} & . & \frac{\partial \bar{x}_3}{\partial x_n} \\ . & . & . & . & . \\ \frac{\partial \bar{x}_n}{\partial x_1} & \frac{\partial \bar{x}_n}{\partial x_2} & \frac{\partial \bar{x}_n}{\partial x_3} & . & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \\ . \\ dx_n \end{pmatrix}$$

Or in tensor notation:

(2.11) $$d\bar{x}^k = \frac{\partial \bar{x}^k}{\partial x^i} dx^i$$

The inverse transformation has the inverse matrix, which follows from the chain rule of differentiating:

(2.12) $$dx^i = \frac{\partial x^i}{\partial \bar{x}^l} d\bar{x}^l$$

and

$$d\bar{x}^k = \frac{\partial \bar{x}^k}{\partial x^i} dx^i = \frac{\partial \bar{x}^k}{\partial x^i}\frac{\partial x^i}{\partial \bar{x}^l} d\bar{x}^l = \frac{\partial \bar{x}^k}{\partial \bar{x}^l} d\bar{x}^l = \delta^{kl} d\bar{x}^l = d\bar{x}^k$$

If we do not have an orthonormal base, which is often the case of generalized coordinates, one can always find a so called *inverse base*.
In the inverse base and in any of the object expanded on the inverse base the indices go upstairs.
Let the base be: $(\vec{e}_1, \vec{e}_2, \vec{e}_3, ..., \vec{e}_n)$ These are linear independent unit vector, which are not (necessarily) mutually orthogonal.
Our aim is now to establish another base: $(\vec{e}^1, \vec{e}^2, \vec{e}^3, ..., \vec{e}^n)$ where the following condition holds:

(2.13) $$\vec{e}_i \cdot \vec{e}^j = \delta_i^j$$

Since $(\vec{e}_1, \vec{e}_2, \vec{e}_3, ..., \vec{e}_n)$ is a base, any of the vectors $(\vec{e}^1, \vec{e}^2, \vec{e}^3, ..., \vec{e}^n)$ can be expanded on that base, which we write it in the following way:

$$\vec{e}^j = \lambda_{1j}\vec{e}_1 + \lambda_{2j}\vec{e}_2 + \lambda_{3j}\vec{e}_3 + ... + \lambda_{nj}\vec{e}_n$$

or

$$\vec{e}^j = \lambda_{kj}\vec{e}_k$$

using the summation convention. The condition (2.13) then gives

$$\vec{e}_i \cdot \vec{e}^j = \lambda_{1j}\vec{e}_i \cdot \vec{e}_1 + \lambda_{2j}\vec{e}_i \cdot \vec{e}_2 + \lambda_{3j}\vec{e}_i \cdot \vec{e}_3 + ... + \lambda_{nj}\vec{e}_i \cdot \vec{e}_n = \delta_i^j$$

We remember that $g_{ij} = \vec{e}_i \cdot \vec{e}_j$ so the system of equations to determine the λ_{ij} is:

(2.14) $$\lambda_{1j}g_{i1} + \lambda_{2j}g_{i2} + \lambda_{3j}g_{i3} + ... + \lambda_{nj}g_{in} = \delta_i^j \qquad j = 1..n$$

If $\underline{\underline{\lambda}}$ is the matrix λ_{ij}, and $\underline{\underline{g}}$ is the matrix g_{jk}, (2.14) can be written:

(2.15) $$\underline{\underline{\lambda}}\,\underline{\underline{g}} = \underline{\underline{E}} \quad \Leftrightarrow \quad \underline{\underline{\lambda}} = \underline{\underline{g}}^{-1}$$

Not surprisingly, the inverse base has the inverse metric form.

From (2.4) $g_{ij} = \vec{e}_i \cdot \vec{e}_j$ it then follows $g^{ij} = \vec{e}^i \cdot \vec{e}^j$

Where $g^{ij} = g_{ij}^{-1}$ is the inverse matrix to g_{ij}
Suppose we have two vectors:

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 + ... + a_n\vec{e}_n \quad \text{and} \quad \vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3 + ... + b_n\vec{e}_n$$

And we form the scalar product

(2.16) $$\vec{a} \cdot \vec{b} = a_i b_j \vec{e}_i \cdot \vec{e}_j = g_{ij} a_i b_j$$

The same vectors expanded on the inverse base give correspondingly:

$$\vec{a}\cdot\vec{b} = a^i b^j \vec{e}^{\,i}\cdot\vec{e}^{\,j} = g^{ij} a^i b^j$$

But if we expand the vector $\vec{a}$ on the base

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 + \ldots + a_n\vec{e}_n$$

and expand the vector $\vec{b}$ on the inverse base

$$\vec{b} = b^1\vec{e}^{\,1} + b^2\vec{e}^{\,2} + b^3\vec{e}^{\,3} + \ldots + b^n\vec{e}^{\,n}$$

we obtain the following expression for the scalar product

(2.17) $$\vec{a}\cdot\vec{b} = a_i b^j \vec{e}_i\cdot\vec{e}^{\,j} = a_i b^j \delta_i^j = a_i b^i \quad \text{so} \quad \vec{a}\cdot\vec{b} = a_i b^i$$

almost as in a Cartesian coordinate system

In the same manner the square of the length of a vector takes its usual form.

(2.18) $$|\vec{a}|^2 = \vec{a}\cdot\vec{a} = a_i a^i$$

3. Tensors

It is required by the laws of physics, that they are covariant or form invariant i.e. "they look the same" under certain coordinate transformations. The laws of mechanics are invariant under translations, rotations and the Galileo transformation. Maxwell's equations are also invariant under the Lorentz transformation of special relativity.
In analytic mechanics one can show, that conservation of momentum is a consequence of invariance to translation, and conservation of angular momentum is a consequence of invariance under rotations.
Thus the invariance of physical objects under various transformations plays an important role.

Mathematical objects which are invariant under certain coordinate transformations are called tensors.
A scalar has only one component, and it is a tensor of rank 0.
An object which has one index (i.e. a vector v_i) is called a tensor of rank 1.
An object which has two indices (i.e. a matrix g_{ij}) is called a tensor of rank 2
Tensors can have arbitrary high rank.

When operating with tensors, one uses consequently the summation convention: that is, if the same index appears twice, summation is implied.

In section 2, we introduced the transformation of differentials from one coordinate system to another. Thus if $\bar{x}_i = \bar{x}_i(x_1, x_2, x_3, \ldots, x_n)$
then

$$d\bar{x}_k = \frac{\partial\bar{x}_k}{\partial x_1}dx_1 + \frac{\partial\bar{x}_k}{\partial x_2}dx_2 + \frac{\partial\bar{x}_k}{\partial x_3}dx_3 + \ldots + \frac{\partial\bar{x}_k}{\partial x_n}dx_n$$

Or in tensor notation, and with the index raised:

(3.1) $$d\bar{x}^k = \frac{\partial \bar{x}^k}{\partial x^i} dx^i$$

Similarly we may write the inverse transformation of $x_i = x_i(\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n)$

(3.2) $$dx^j = \frac{\partial x^j}{\partial \bar{x}^k} d\bar{x}^k$$

A scalar has the same value in all coordinate systems: $\varphi(\bar{x}) = \varphi(x)$

Any object a^i that transform as the coordinate differentials is called a *contravariant* tensor of rank 1.

(3.3) $$\bar{a}^k = \frac{\partial \bar{x}^k}{\partial x^i} a^i$$

On the other hand the gradient of a scalar, transforms as:

$$\frac{\partial \varphi}{\partial \bar{x}^k} = \frac{\partial \varphi}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^k}$$

Any object b_i that transform as the gradient of a scalar is called a *covariant* tensor of rank 1.

(3.4) $$\bar{b}_k = b_j \frac{\partial x^j}{\partial \bar{x}^k}$$

A covariant tensor c_{ij} has the transformation properties:

(3.5) $$\bar{c}_{ij} = c_{kl} \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial x^l}{\partial \bar{x}^j}$$

A mixed tensor is a tensor with having index both upstairs and downstairs.

Notice, with the transformation properties defined above, the scalar product is an invariant.

(3.6) $$\bar{\vec{a}} \cdot \bar{\vec{b}} = \bar{a}_i \bar{b}^i = a_j \frac{\partial x^j}{\partial \bar{x}^i} \frac{\partial \bar{x}^i}{\partial x^l} b^l = \delta_l^j a_j b^l = a_j b^j$$

With the help of the metric tensor g_{ij} or its inverse tensor g^{ij} you may raise or lower one or several indices. This comes about, because the metric tensor is the transformation matrix between a base of unit vectors and the inverse base.

$$a_i g^{ij} = a^j \qquad\qquad c^{ij} g_{ik} g_{jl} = c_{jl}$$

For the same reason, you may also *contract* two tensors, by making equal a lower index in one with the upper index in the other.

$$c_i = a_{i\,j} b^j$$

That c_i actually has the proper transformation properties of a covariant vector, can be shown by writing the transformation properties for $a_{i\,j}$ and b^j, as we did for the scalar product.

4. Covariant derivative

We have seen above, that for an object to be a tensor, it must have certain transformation properties. Since generalized coordinates require a base and its inverse base, which are orthogonal to each other, we have covariant and contravariant tensors, distinguished from each other by having lower and upper indices.

A contravariant vector transforms in the same way, as the coordinate transformation: $d\bar{x}^k = \frac{\partial \bar{x}^k}{\partial x^i} dx^i$

Any object a^i that transform as the coordinate differentials is called a *contravariant* tensor of rank 1.

$$\bar{a}^k = \frac{\partial \bar{x}^k}{\partial x^i} a^i \qquad (4.1)$$

On the other hand the gradient of a scalar, transforms like

$$\frac{\partial \varphi}{\partial \bar{x}^k} = \frac{\partial \varphi}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^k}$$

Any object b_i that transform as the gradient of a scalar is called a *covariant* tensor of rank 1.

$$\bar{b}_k = b_j \frac{\partial x^j}{\partial \bar{x}^k} \qquad (4.2)$$

One might expect that the derivatives of the components of a vector would transform like a tensor, but that is not the case. This is one of the reasons for introducing the concept of *covariant derivative*, a derivative which has proper tensor properties.
The transformation of the i'th component of a contravariant vector v^i goes as follows.
According to (4.2)

$$\bar{v}^i = \frac{\partial x^i}{\partial \bar{x}^k} v^k$$

Consistently

$$\frac{\partial \bar{v}^i}{\partial \bar{x}^j} = \frac{\partial}{\partial \bar{x}^j}\left(\frac{\partial \bar{x}^i}{\partial x^k} v^k\right) = \frac{\partial^2 \bar{x}^i}{\partial \bar{x}^j \partial x^k} v^k + \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial v^k}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^j}$$

$$\frac{\partial \bar{v}^i}{\partial \bar{x}^j} = \frac{\partial}{\partial \bar{x}^j}\left(\frac{\partial \bar{x}^i}{\partial x^k} v^k\right) = \frac{\partial^2 \bar{x}^i}{\partial \bar{x}^j \partial x^k} v^k + \left(\frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial x^l}{\partial \bar{x}^j}\right) \frac{\partial v^k}{\partial x^l}$$

If only the last term was present, the derivative of v^i would transform as a proper mixed tensor, but the first term prevents this.

The reason for this, is that the components of a vector are the projection on a base vectors, and they are position dependent, $\vec{e}^i = \vec{e}^i(x)$ and $v^i = \vec{v} \cdot \vec{e}^i$.
Differentiating $v^i = \vec{v} \cdot \vec{e}^i$, we get:

(4.3) $$\frac{\partial v^j}{\partial x^i} = \frac{\partial}{\partial x^j}(\vec{v} \cdot \vec{e}^j) = \frac{\partial \vec{v}}{\partial x^i} \cdot \vec{e}^j + \vec{v} \cdot \frac{\partial \vec{e}^j}{\partial x^i}$$

It is the second term that causes the trouble, and has to be dealt with.

The idea is now to try to construct a differential operator D_i that has tensor properties, when applied on components of a vector v^j. That is:

(4.4) $$\overline{D}_i \overline{v}^j = \frac{\partial x^k}{\partial \overline{x}^i} \frac{\partial \overline{x}^j}{\partial x^l} D_k v^l$$

Any vector $\overline{v}$ can be expanded either on the base $\vec{e}_i$ or on the inverse base $\vec{e}^j$.
The problem arises, because the unit vectors themselves depend on position, which is not the case in a Cartesian coordinate system.
If we expand $\vec{v}$ on the base $\vec{e}_i$: $\vec{v} = v^i \vec{e}_i$ then the projections are $v^j = \vec{v} \cdot \vec{e}^j$ since $\vec{e}_i \cdot \vec{e}^j = \delta_i^j$.
The point is that both $\overline{v}$ and $\vec{e}^j$ are good tensors, so when we differentiate v^j written as $\vec{v} \cdot \vec{e}^j$ it might just be the differential operator we are looking for, since

$$\frac{\partial \vec{v}}{\partial \overline{x}^i} = \frac{\partial x^j}{\partial \overline{x}^i} \frac{\partial \vec{v}}{\partial x^j} \qquad \text{and} \qquad \vec{\overline{e}}^j = \frac{\partial \overline{x}^j}{\partial x^k} \vec{e}^k$$

By taking the scalar product of the two vectors we then obtain

$$\frac{\partial \vec{v}}{\partial \overline{x}^i} \cdot \vec{\overline{e}}^j = \left(\frac{\partial x^j}{\partial \overline{x}^i} \frac{\partial \overline{x}^j}{\partial x^k}\right) \frac{\partial \vec{v}}{\partial x^j} \vec{e}^k$$

Which shows that it transforms like a proper mixed tensor, and it might be the covariant derivative D_i that we are looking for.

(4.5) $$D_i v^j = \frac{\partial \vec{v}}{\partial x^i} \cdot \vec{e}^j$$

This relation implies that $D_i v^j$ can be viewed as projections of the vectors $\frac{\partial \vec{v}}{\partial x^i}$ on the vectors $\vec{e}^j$,

we may therefore also write $\frac{\partial \vec{v}}{\partial x^i}$ expanded on the basis vectors $\vec{e}_k$

(4.6) $$\frac{\partial \vec{v}}{\partial x^i} = D_i v^k \vec{e}_k$$

Taking the scalar product with $\vec{e}^{\,j}$ we get:

(4.6) $$\frac{\partial \vec{v}}{\partial x^i}\cdot \vec{e}^{\,j} = D_i v^k \vec{e}_k \cdot \vec{e}^{\,j} = D_i v^k \delta_k^j = D_i v^j$$

which is (4.5).

4.1 Christoffel symbols as expansion coefficients

However we do not have a similar simple transformation property for the base vectors $\vec{e}_k$, since they by definition change under coordinate transformations.
As a result the corresponding expansion for (4.6) $\frac{\partial \vec{v}}{\partial x^i} = D_i v^k \vec{e}_k$ can not be copied for $\frac{\partial \vec{e}^{\,j}}{\partial x^i}$,
but instead the derivative is written symbolically as an expansion on the vectors $\vec{e}^{\,j}$

(4.7) $$\frac{\partial \vec{e}^{\,k}}{\partial x^i} = -\Gamma^k_{i\,j}\vec{e}^{\,j}$$

Or when forming the scalar product with $\vec{v}$

$$\vec{v}\cdot\frac{\partial \vec{e}^{\,k}}{\partial x^i} = -\Gamma^k_{i\,j} v^j$$

Putting $\vec{v}\cdot\frac{\partial \vec{e}^{\,k}}{\partial x^i} = -\Gamma^k_{i\,j} v^j$ and $D_i v^j = \frac{\partial \vec{v}}{\partial x^i}\cdot\vec{e}^{\,j}$ into $\frac{\partial v^j}{\partial x^i} = \frac{\partial \vec{v}}{\partial x^i}\cdot\vec{e}^{\,j} + \vec{v}\cdot\frac{\partial \vec{e}^{\,j}}{\partial x^i}$ we get:

(4.8) $$\frac{\partial v^j}{\partial x^i} = \frac{\partial \vec{v}}{\partial x^i}\cdot\vec{e}^{\,j} + \vec{v}\cdot\frac{\partial \vec{e}^{\,j}}{\partial x^i} = D_i v^j - \Gamma^k_{i\,j} v^j$$

And after isolating $D_i v^j$, we finally arrive at.

(4.9) $$D_i v^j = \frac{\partial v^j}{\partial x^i} + \Gamma^j_{i\,k} v^k$$

Thus in order to produce a derivative of a vector component that conforms to the transformation properties of a tensor, we must add an extra term. Neither of the two terms on the right side of (4.9) are tensors, but the sum of them is.
As we shall see later the concept of covariant derivative has affinities to finding the shortest path on a between to points on a manifold, and to the concept of parallel displacement.

One can show that de covariant derivative of a covariant vector is given by:

(4.10) $$D_i v_j = \frac{\partial v^j}{\partial x^i} - \Gamma^k_{i\,j} v^k$$

4.2 Expressing the Christoffel symbol by the metric tensor

You may find the covariant derivative for at mixed T^i_j tensor as the covariant derivative of the direct product $v^i u_j$. It is differentiated covariantly, treating each index separately. For a mixed tensor the result is.

(4.11) $$D_i T^k_j = \frac{\partial T^k_j}{\partial x^i} - \Gamma^l_{ij} T^k_l + \Gamma^k_{im} T^m_j$$

A specific example is the differentiating of the metric tensor g_{ij}

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

The Christoffel symbol Γ^k_{ij} is symmetric when permuting the two lower indices.
This follows from the expression (4.17) below, where the Christoffel symbol is expressed by means of the derivatives of the metric tensor.

We shall now prove that although the metric tensor is position dependent, it is constant with respect to covariant differentiation. $D_k g_{ij} = 0$ (Identically zero).

One may prove this, by using the definition of the metric tensor in terms of the base vectors:

$$g_{ij} = \vec{e}_i \cdot \vec{e}_j$$

$$\frac{\partial g_{ij}}{\partial x^k} = \frac{\partial}{\partial x^k}(\vec{e}_i \cdot \vec{e}_j) = \frac{\partial \vec{e}_i}{\partial x^k} \cdot \vec{e}_j + \vec{e}_i \cdot \frac{\partial \vec{e}_j}{\partial x^k}$$

$$= (\Gamma^l_{ki} \vec{e}_l) \cdot \vec{e}_j + (\Gamma^l_{kj} \vec{e}_l) \cdot \vec{e}_i$$

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{li} \qquad \Rightarrow$$

(4.13) $$\frac{\partial g_{ij}}{\partial x^k} - \Gamma^l_{ki} g_{lj} - \Gamma^m_{kj} g_{mi} = 0 \qquad \Leftrightarrow \qquad D_k g_{ij} = 0$$

We are now ready to express the Christoffel symbols, by the derivatives of the metric tensor.
The starting point is (4.13)

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

Rewriting (4.13) cyclically permuting the indices *i, j, k*, and using the symmetry properties of the Christoffel symbol $\Gamma^k_{ij} = \Gamma^k_{ji}$ and the same symmetry of the metric form $g_{ij} = g_{ji}$ we get.

(4.14) $$\frac{\partial g_{ki}}{\partial x^j} - \Gamma^l_{jk} g_{li} - \Gamma^l_{ji} g_{kl} = 0$$

(4.15) $$\frac{\partial g_{jk}}{\partial x^i} - \Gamma^l_{ij} g_{kl} - \Gamma^l_{ik} g_{jl} = 0$$

Adding the two first equations, and subtracting the last (after having permuted some indices), we arrive at:

$$\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i} - 2\Gamma^l_{kj} g_{il} = 0$$

(4.16) $$\Gamma^l_{kj} g_{il} = \tfrac{1}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)$$

Finally by multiplying both sides by the inverse matrix of the metric form g^{lm}, we get

$$\Gamma^l_{kj} g_{il} g^{lm} = \tfrac{1}{2} g^{lm}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)$$

$$\Gamma^l_{kj} \delta^m_i = \tfrac{1}{2} g^{lm}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)$$

$$\Gamma^m_{kj} = \tfrac{1}{2} g^{lm}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)$$

Or by renaming some of the indices

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

4.3 Geodesics and covariant derivatives

In its origin, a geodesic is the name of the shortest path between two positions on earth. As we know that for a sphere it is a great circle connecting the two positions.
In Differential geometry a geodesic is the shortest path connecting two points on a manifold.

In Differential Geometry 1, we restricted ourselves to surfaces i.e. two dimensional manifolds. But here we shall show that the results also apply to higher dimensions.

We shall do so by finding differential equations for the geodesic parameter curve.

We assume that $x = (x^1(t), x^2(t), x^3(t), ..., x^n(t))$ is a parameter curve with the parameter *t*.
The distance element is:

(4.18) $$ds = \sqrt{g_{ij} dx^i dx^j}$$

and therefore

(4.19) $$\dot{s} = \frac{ds}{dt} = \sqrt{g_{ij} \dot{x}^i \dot{x}^j}$$

Where, as usual a bullet above a variable means differentiation with respect to *t*.

To derive the differential equations for $x^i(t)$, which give the minimum value for *s* we use the classical Lagrangian approach.

$$\delta \int L dt = 0$$

where $L = L(x, y, \dot{y}, t)$ and the solution is the Euler-Lagrange equation.

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

If there are several variables $x = (x^1(t), x^2(t), x^3(t), ..., x^n(t))$ the equations become:

(4.20) $$\frac{\partial L}{\partial x^k} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^k} = 0 \qquad \text{with} \qquad L = \sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)}$$

Note that $g_{ij}(x^i, x^j)$ does not depend explicitly on *t*.

If we furthermore choose *t* as a *natural parameter,* we have

$$\frac{ds}{dt} = \frac{ds}{ds} = \sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)} = 1$$

We may then drop the square root in the calculations, i.e.

$$\frac{\partial}{\partial x^k} \sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)} =$$

$$\frac{1}{2\sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)}} \frac{\partial}{\partial x^k} g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t) = \tfrac{1}{2} \frac{\partial}{\partial x^k} (g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t))$$

And by the same token, when differentiating with respect to *t*. Thus we find:

$$\frac{\partial L}{\partial x^k} = \frac{\partial}{\partial x^k} \sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)} = \tfrac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^i \dot{x}^j$$

and

$$\frac{\partial L}{\partial \dot{x}^k} = \frac{\partial}{\partial \dot{x}^k} \sqrt{g_{ij}(x^i, x^j)\, \dot{x}^i(t) \dot{x}^j(t)} = \tfrac{1}{2}(g_{kj} \dot{x}^j + g_{ik} \dot{x}^i) \;\; = g_{ik} \dot{x}^i \qquad (\text{Since } g_{ik} = g_{ki})$$

Multiplying by 2, the Euler-Lagrange equation becomes

$$\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j - 2\frac{d}{dt}g_{ik}\dot{x}^i = 0$$

$$\frac{d}{dt}g_{ik}\dot{x}^i = \frac{\partial g_{ik}}{\partial x^j}\dot{x}^j\dot{x}^i + g_{ik}\ddot{x}^i$$

And, because of the symmetry of the metric form: $\frac{\partial g_{ik}}{\partial x^j} = \frac{1}{2}(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i})$

Putting it all together we get

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} = 0 \qquad \Rightarrow$$

$$\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j - (\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i})\dot{x}^i\dot{x}^j - 2g_{ik}\ddot{x}^i = 0$$

Multiplying by ½ , changing sign and rearranging the terms we finally get:

$$g_{ik}\ddot{x}^i + \frac{1}{2}(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k})\dot{x}^i\dot{x}^j = 0$$

Finally multiplying by the inverse matrix to g_{ik} , g^{kl} we arrive at the equation.

$$g^{kl}g_{ik}\ddot{x}^i + \frac{1}{2}g^{kl}(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k})\dot{x}^i\dot{x}^j = 0 \qquad \Leftrightarrow$$

$$\delta_{il}\ddot{x}^i + \frac{1}{2}g^{kl}(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k})\dot{x}^i\dot{x}^j = 0 \qquad \Leftrightarrow$$

(4.21) $$\ddot{x}^l + \Gamma^l_{ij}\dot{x}^i\dot{x}^j = 0$$

Or, if renaming *l* to *k*

(4.21) $$\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$$

This we shall compare to the covariant derivative

$$D_i v^j = \frac{\partial v^j}{\partial x^i} + \Gamma^j_{ik}v^k$$

put to 0.

We choose the vector $v^{j,}$ to be the tangent vector $\frac{dx^j}{ds}$ of a curve

$$x = x(s) = (x^1(s), x^2(s), x^3(s), \ldots, x^n(s))$$

We then change the formula for the covariant derivative, by differentiating down to s, remembering, that when we differentiate with respect to x^j, should be followed by multiplication with $\frac{dx^j}{ds}$. The covariant derivative now writes, since it is the covariant derivative with respect to x^j.

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

Insertion $v^j = \frac{dx^j}{ds}$ one finds:

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

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

Comparing with the equation for the geodesic

$$\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$$

We see that it is in fact the same equation, apart from the name of the indices..
If you move in 3-dim space along a straight line, i.e. in the direction of a constant vector, the derivative of that vector is zero, at any point.

This suggests that, when you move in a manifold along a curve with zero covariant derivative, you move along a "straight line" meaning the shortest path between any two points on the curve.
The geodesics in space-time are the paths followed by a light ray, since it is the geodesics for a body falling freely in the influence of gravity.

4.4 Directional change under parallel transport of a vector

Parallel transport is a fundamental notion in differential geometry. It comes about because the change of a vector has two components.

One is the dependence of the vector on coordinates: $\frac{\partial v^j}{\partial x^i}$

The second is the coordinate themselves, which change with position, represented by the basic vectors $\vec{e}^k = \vec{e}^k(x)$ where $x = (x^1, x^2, x^3, .., x^n)$.

We have already seen this, because this is precisely, what is reflected in the difference between the normal derivative, which only includes the first part, and the covariant derive which includes both parts. As previously covariant differentiation with respect to x^i is denoted by a capital D_i.

According to (4.9) $$D_i v^j = \frac{\partial v^j}{\partial x^i} + \Gamma^j_{ik} v^k$$

Or when expressing it with total differentials:

$$Dv^j = (D_i v^j)dx^i \qquad \text{And} \qquad dv^j = \frac{\partial v^j}{\partial x^i} dx^i$$

So

$$dv^j = \frac{\partial v^j}{\partial x^i} dx^i = (D_i v^j)dx^i - \Gamma^j_{ik} v^k dx^i \qquad \text{or}$$
$$dv^j = d(\vec{e}^j \cdot \vec{v}) = \vec{e}^j \cdot d\vec{v} + \vec{v} \cdot d\vec{e}^j = Dv^j - \Gamma^j_{ik} v^k dx^i$$

The notion of parallel transport comes about, when moving a vector without changing the vector itself, that is, $\vec{e}^j \cdot d\vec{v} = Dv^j = 0$.
In a Euclidian space the notion of parallel transport is the same as parallel displacement: that is, the vector is unchanged.

Setting the "true change", to zero

(4.23) $$Dv^j = 0 \quad \Leftrightarrow \quad dv^j + \Gamma^j_{ik} v^k dx^i = 0 \qquad \text{(True change)}$$

We recall, that $D_i g_{ij} = 0$, which reflects the fact that the change of the metric tensor is due to coordinate changes only. This is however the same, as the covariant derivative is identically zero along a specific curve.
The process of parallel transporting a vector v^j along a parameter curve $x^i(t)$, can according to (4.12) be accomplished as:

(4.24) $$\frac{Dv^j}{Dt} = \frac{dv^j}{dt} + \Gamma^j_{ik} v^k \frac{dx^i}{dt} = 0$$

In this manner, we may define the straightest possible curve by the condition of it being the "line" constructed by parallel transport of its tangent vector.

In this way the condition can also be formulated setting $v^j = \dfrac{dx^j}{dt}$, so according to (4.24)

(4.25) $$\frac{D}{Dt}\left(\frac{dx^j}{dt}\right) = \frac{d^2x^j}{dt^2} + \Gamma^j_{ik} v^k \frac{dx^i}{dt}\frac{dx^j}{dt} = 0$$

An equation we (again) recognize as the geodesic equation.
Covariant derivative, parallel transport, geodesics are basically the same geometrical concept.
We may then conclude that when a vector v^i is parallel transported along a geodesic, one can show that the angle subtended by the vector and the geodesic: that is the, tangent of the geodesic, is unchanged.

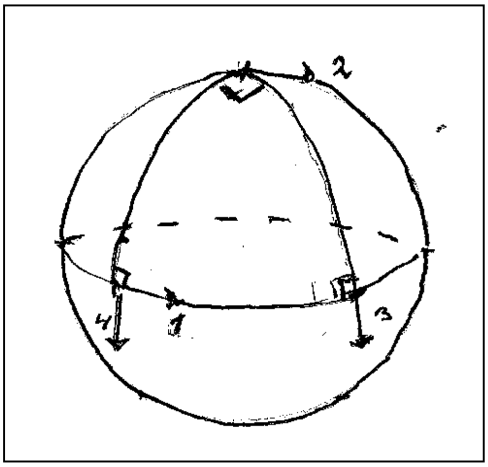
4.5 Angular excess and its correspondence to curvature

Curvature measures the amount of deviation the manifold differs from Euclidian space.
In a two dimensional space one can argue that for an infinitesimal polygon (a triangle) there is a remarkably simple relation between the angular excess E (which is the difference between the sum of the angles in the polygon and the sum of angles in the corresponding Euclidian polygon) and the area dA of the polygon. The relation is proportionality, and the scaling factor is the Gaussian curvature K.

(4.27) $$E = K\, dA$$

How the angular excess can be determined is not quite obvious, but we shall use the concept of parallel transport to illustrate the method. This has the advantage that it will allow us to generalize the relation (4.27) from a two dimensional surface into n-dimensional manifolds.

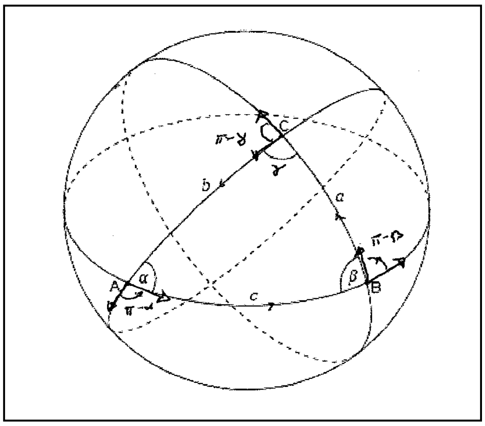
It can be shown that the angular excess of a polygon E is equal to the directional change of a vector when parallel transported around the polygon.
We shall first illustrate this by a simple example, where it is immediately clear.
See the figure below.



The figure shows a sphere, and a spherical triangle with three 90^0 angles. A vector initially placed at (1) is parallel transported to (2). Since the vector is perpendicular to the great circle arc 1 - 2, it does not change direction. But when it reaches (2) it is tangent to the arc 2 -3.
When parallel transported: that is along the arc, the directional change is 90^0.
When transported from (3) to (4), again there is no directional change, as the vector is perpendicular to the arc (3) – (4).
We see that the directional change of the vector, when parallel transported from (1) to (4) is 90^0, which is equal to the angular excess of the triangle.

For the directional change of a parallel transportation of a vector along an arbitrary triangle one may argue as follows: See the figure below.

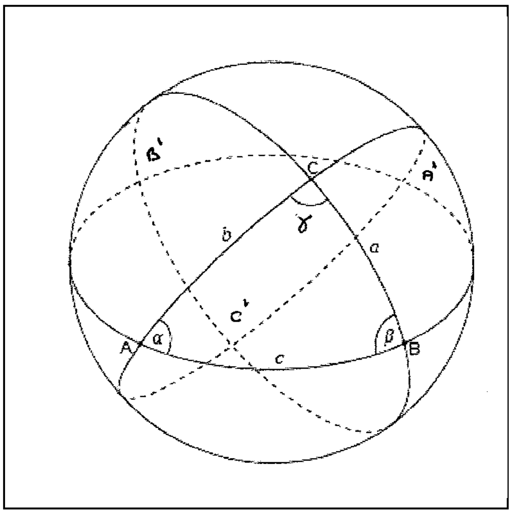
In the case for a constant curvature i.e. a sphere, the area does not need to be infinitesimal, and we shall now prove (4.27) for a spherical triangle.



The figure shows a spherical triangle *ABC*, with the three angles α, β, γ. We notice, that it is only the part of a vector along a curve, which has a directional change.
We start out at *A*, and move the tangent vector by parallel transport to *B*. The directional change is d_1.
Then we turn the vector an angle π - β, so the vector now is along *a*. We move the tangent vector by parallel transport to *C*, and the vector gets a directional change d_2.
Then we turn the vector an angle π - γ, so it becomes tangent to the arc *AC* and move the tangent vector by parallel transport to *A*.The directional change is d_3.
Finallly we turn the vector an angle π – α, so it coincides with the initial vector. The net result is that the vector has been turned an angle
π – β + π – γ + π – α = π – (α +β + γ).
At the same time, it has undergone a directional change $d_1 + d_2 + d_3$.
By construction the sum of the angular change, and the directional change must be zero.

(4.28) $$d_1 + d_2 + d_3 + \pi - (\alpha + \beta + \gamma) = 0 \quad \Leftrightarrow \quad d_1 + d_2 + d_3 = \alpha + \beta + \gamma - \pi$$

Which illustrates the assertion that when a vector is transported parallel around a polygon, the directional change is equal to the angular excess.



If you look at the figure, one can see, that there are six lunes.
ABA'C, AC'A'B', BCB'A, BC'B'A', CAC'B, CA'C'B'.
Each point on the sphere belongs to exactly one of the six lunes, except for the two triangles *ABC* and *A'B'C',* which belongs to three lunes.
The area of a lune, with an opening angle θ is the fraction of θ to the perimeter 2π*R* times the area of the sphere 4πR^2
Area of lune $= \dfrac{\theta}{2\pi} 4\pi R^2 = 2\theta R^2$.
The area of the six lunes is therefore: $2(\alpha + \beta + \gamma)R^2$.
This is the equal to the area of the sphere, minus two times the area of the triangles *ABC* an *A'B'C'*.
For spherical triangles, however congruity of angles means congruity of the triangles themselves.
Let A_{abc} be the common area of the two triangles, we can there write an equation for the areas.

$$4(\alpha + \beta + \gamma)R^2 = 4\pi R^2 - 4A_{ABC}$$

By divding by 4R^2, we obtain:

(4.29) $$(\alpha + \beta + \gamma) - \pi = \frac{A_{ABC}}{R^2}$$

Comparing this to (4.27) Angular excess = Curvature · Area:

(4.29) $$E = K \cdot A,$$

we see, that it gives the Gaussian curvature R^{-2}, which is in fact the case for a sphere.

For a sphere, the result (4.29) can easily be extended to an arbitrary spherical polygon, since such a polygon
can always be divided in triangles, where the sum of angular excesses equals the angular excess of the polygon, and the sum of the areas of the triangles equals the area of the polygon.

Also we have seen in (4.28) that the angular excess E of a polygon is equal to the directional change Δv of a vector, when it is parallel transported along the edges of the polygon.
The angular change of a vector is $\frac{dv}{v}$ therefore (4.29) can also be written:

(4.30) $$\frac{dv}{v} = KdA \quad \Leftrightarrow \quad dv = KvdA$$

The above theorem is only valid, when the curvature is constant, and therefore it applies in generel only to infinitesimal polygons.

However when we no longer consider only 2-dimensional surfaces, and want to generalize to higher dimensions the concept of area, becomes a bit more circomstantial.

In 3-dimensional Euclidian space, the area spanned by two vectors ***a*** and ***b*** , can be calculated as the length of the cross product $\vec{\sigma} = \vec{a} \times \vec{b}$. Or using the Levi-Civitas antisymmetric tensor ε_{ijk} .

(4.30) $$\sigma_k = \varepsilon_{ijk} a^i b^j$$

Where $\vec{\sigma}$ has the magnitude $|\vec{a}||\vec{b}|\sin\theta$.
(4.30) cannot immediately be generalized to higher dimensions in this form.
For higher dimensions we shall need an antisymmetric tensor with a different number of indices.
In four dimensions the Levi Civitas symbol becomes ε_{ijkl} and so on.

In four dimensions we shall use a two index object σ^{ij} to represent the area.

(4.31) $$\sigma^{ij} = \varepsilon^{ijk}\sigma_k = \varepsilon^{ijk}\varepsilon_{mnk} a^m b^n = \tfrac{1}{2}(a^i b^j - a^j b^i),$$

Where we have made use of the identity: $\varepsilon^{ijk}\varepsilon_{mnk} = \tfrac{1}{2}(\delta^i_m \delta^j_n - \delta^i_n \delta^j_m)$
For this relation the index 1..3 is actually irrelevant, and we can readily generalize (4.31) to n dimensions: σ^{ij} , where i, j = 1,2,3…,n. So in any case, the area element is represented by:

(4.32) $$\sigma^{ij} = \tfrac{1}{2}(a^i b^j - a^j b^i)$$

4.7 Generalization of the Riemann tensor to n-dimensional space

The equation (4.30) and (4.32): that is, the tensor expression for the area, suggest that one may express the change of a vector due to a parallel transport around a parallelogram spanned by the vectors a^k and b^l where the curvature K is replaced by the Riemann tensor.

(4.33) $$dv^i = R^i_{jkl} v^j d\sigma^{kl}$$

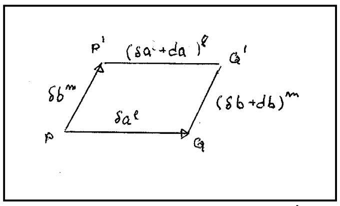
Since all other parts of the equation are good tensors, it follows from the quotient theorem that R^i_{jkl} is also a tensor.

By the rather intriguing calculation below (I suggest that you skip it), one can show that:

(4.34) $$R^i_{jkl} = \frac{\partial \Gamma^i_{jl}}{\partial x^k} - \frac{\partial \Gamma^i_{jk}}{\partial x^l} + \Gamma^i_{mk}\Gamma^m_{jl} - \Gamma^i_{ml}\Gamma^m_{jk}$$

Since the Christoffel symbol being first derivative of the metric tensor, the Riemann curvature tensor is a nonlinear function of $\frac{\partial^2 g}{(\partial x^k)^2}$ *and* $(\frac{\partial g}{\partial x^k})^2$

Derivation of (4.34)



The figure shows a parallelogram *PP'Q'Q* spanned by the infinitesimal vectors δa^l *and* δb^m. The opposite vectors $(\delta a + da)^l$ *and* $(\delta b + db)^m$, are obtained by parallel transport of δa^l *and* δb^m respectively.
From the figure we see that $db^l = \delta a^l$ *and* $da^m = \delta b^m$
Recall that parallel transport of a vector v^j implies that the covariant derivative is zero: $D_i v^j = 0$, meaning that the total change of the vector is due to coordinate changes only.

(4.35) $$dv^i = -\Gamma^i_{jk} v^j dx^k$$

Our aim is to derive the expression above (4.34) from parallel transporting a vector around a closed path, given by the parallelogram PP'Q'Q. From (4.35) we then get:

(4.36) $$(\delta a + da)^l = \delta a^l - \Gamma^l_{ij}\delta a^i \delta b^j \quad \text{and} \quad (\delta b + db)^m = \delta a^m - \Gamma^m_{ij}\delta a^i \delta b^j$$

We then calculate from (4.35) the change of a vector v^i due to parallel transport from *P* ->*Q* -> *Q'*.

(4.36) $$\begin{aligned} dv^i{}_{PQQ'} &= dv^i{}_{PQ} + dv^i{}_{QQ'} \\ &= -(\Gamma^i_{jl} v^j)_P \delta a^l - (\Gamma^i_{jm} v^j)_Q (\delta b + db)^m \end{aligned}$$

Where *P* and *Q* denote the respective positions, where these functions are to be evaluated.
Since the aim is to evaluate and compare all quantities at the same position *P*, we shall make a Taylor expansion around *P* for the quantities evaluated at *Q*.

(4.37) $$(\Gamma^i_{jm})_Q = (\Gamma^i_{jm})_P + (\frac{\partial \Gamma^i_{jm}}{\partial x^l})_P \delta a^l$$

And according to (4.35)

$$(v^i)_Q = (v^i)_P + (\frac{\partial v^j}{\partial x^l})_P \delta a^l = (v^i)_P - (\Gamma^i_{kl} v^k)_P \delta a^l$$

Substituting (4.36) and (4.37) into (4.36), and dropping the index *P*, since all quantities, are now evaluated at *P*.

$$dv^i{}_{PQQ'} = -\Gamma^i_{jl} v^j \delta a^l - (\Gamma^i_{jm} + \frac{\partial \Gamma^i_{jm}}{\partial x^l} \delta a^l)(v^j - \Gamma^j_{kl} v^k \delta a^l)(\delta b^m - \Gamma^m_{pq} \delta a^p \delta b^q)$$

Multiplying out the brackets, keeping only terms up to $\delta a^p \delta b^q$ we find:

(4.38) $$dv^i{}_{PQQ'} = -\Gamma^i_{jl} v^j \delta a^l - \Gamma^i_{jm} v^j \delta b^m + \Gamma^i_{jm}\Gamma^m_{pq} v^j \delta a^p \delta b^q - \frac{\partial \Gamma^i_{jm}}{\partial x^l} v^j \delta a^l \delta b^m + \Gamma^i_{jm}\Gamma^j_{kl} v^k \delta a^l \delta b^m$$

The directional change by the path *PP'Q'* can simply be obtained from the above expression by interchanging *a* with *b*.

(4.39) $$dv^i{}_{PQP'} = -\Gamma^i_{jl} v^j \delta b^l - \Gamma^i_{jm} v^j \delta a^m + \Gamma^i_{jm}\Gamma^m_{pq} v^j \delta b^p \delta a^q - \frac{\partial \Gamma^i_{jm}}{\partial x^l} v^j \delta b^l \delta a^m + \Gamma^i_{jm}\Gamma^j_{kl} v^j \delta a^l \delta b^m$$

To find the collected directional change from a tour around the parallelogram PQP'Q' we subtract (4.39) from (4.38).

(4.40) $$dv^i = dv^i{}_{PP'Q'} - dv^i{}_{PQQ'} = (\frac{\partial \Gamma^i_{km}}{\partial x^l} - \frac{\partial \Gamma^i_{kl}}{\partial x^m} + \Gamma^i_{jl}\Gamma^l_{km} - \Gamma^i_{jm}\Gamma^j_{kl}) v^k \delta a^l \delta b^m$$

Because the combination in the bracket is antisymmetric with respect to the indices *l* an *m*, only the antisymmetric combination ½($a^l b^m - a^m b^l$) contributes, but this is just the area tensor, we introduced in (4.31)

Comparering with (4.33) with a minor change in the name of the indices, we realize, that (4.40) is identical with (4.31). And we have found an expression for the Riemann tensor.

$$dv^i = R^i_{jkl} v^j d\sigma^{kl}$$

4.8 Symmetries and contractions of the Riemann curvature tensor

The Riemann curvature tensor has certain symmetry properties reducing the number of independent components in 4-dimensional space from 64 to 20.

On a surface in a 2-dimensional space, the Riemann tensor has only one component which is the Gaussian curvature *K*.

One can lower the upper index of the Riemann tensor by applying the metric tensor g_{ij}.

(4.41) $$R_{ijkl} = g_{im} R^m_{jkl}$$

The Riemann tensor is anti symmetric with respect to interchanging the first and the second index, and that of the third and the fourth index, respectively.

(4.42) $$R_{ijkl} = -R_{jikl} \quad and \quad R_{ijkl} = -R_{ijlk}$$

It is symmetric with respect to interchange of the pair made up of the first and the second indices, and the pair made up of the third and the fourth indices.

(4.43) $$R_{ijkl} = R_{klji}$$

It also has a cyclic symmetry

(4.44) $$R_{ijkl} + R_{iljk} + R_{ikjl} = 0$$

Because of the symmetry properties stated above the contractions of the curvature tensor are essentially unique. After demonstrating the various contractions we show how we arrive at Einstein's tensor, which is the left part of field equation in General Relativity.

The Ricci tensor is the Riemann curvature tensor, with the first and the third indices contracted-

(4.45) $$R_{ij} = g^{lm} R_{limj} = R^{m}_{imj}$$

Which is a symmetric tensor:

$$R_{ij} = R_{ji}$$

The Ricci scalar is the Riemann curvature tensor contracted twice.

(4.46) $$R = g^{ij} R_{ij}$$

The Bianchi identities are a set of constraints on the curvature tensor

(4.47) $$D_k R_{ijlm} - D_j R_{kilm} + D_i R_{jklm} = 0$$

We shall not prove these assertions, but rather perform some contractions with the metric tensor. Since the metric tensor is covariantly constant. $D_k g^{lm} = 0$, the contraction can be "pushed through" the covariant differentiation.

(4.48) $$D_k R - D_j g^{jm} R_{km} - D_i g^{il} R_{kl} = 0$$

At the last two terms, the contraction just raises the index.

$$D_k R - D_j R_k^j - D_i R_k^i = D_k R - 2 D_j R_k^j = 0$$

Pushing yet another g^{ik} to raise the index in the last term, and putting the covariant derivative outside, we find:

(4.49) $$D_k (R g^{ik} - 2R^{ik}) = 0$$

We see, that the combination

(4.50) $$G^{ij} = R^{ij} - \tfrac{1}{2} R g^{ij}$$

Is covariantly constant, i.e. its covariant derivative is identically zero.

$$D_i G^{ij} = 0$$

The reason for developing the last section is that G^{ij} is the tensor that appears on left side of Einstein's famous field equation in The General Theory of Relativity, and for that reason, it is called the Einstein tensor.

To elaborate on this, is beyond a text on differential geometry and tensor analysis, here we shall just state Einstein's geometrical interpretation of the dynamics of moving in a gravitational field.

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

The geodesic of a particle falling freely in a gravitational field is proportional to the energy momentum tensor.

If one takes the Newtonian limit of (4.51) the proportional factor κ can be shown to be:

(4.52) $$\kappa = -\frac{8\pi G}{c^4}$$

Where G is the same constant, as in Newton's law of gravitation, where F is the attracting force between two masses m_1 and m_2 , when separated by a distance r.

$$F = G\frac{m_1 m_2}{r^2}$$

References:
Tai-Pei Cheng: Relativity, Gravitation and Cosmology.
Jon Mathews and Robert L. Walker: Mathematical methods of physics.
Albert Einstein: The meaning of relativity