

Tensors, and differential forms - Lecture 2

1 Introduction

The concept of a tensor is derived from considering the properties of a function under a transformation of the coordinate system. A description of a physical process cannot depend on the coordinate orientation or origin, and this principle can be expanded to look for other mathematical symmetries. To begin it is best to start the description of tensors by reviewing the transformation of a vector function under a rotation of coordinates. Consider the vector function;

$$\vec{F}(x, y, z) = f_x(x, y, z)\hat{x} + f_y(x, y, z)\hat{y} + f_z(x, y, z)\hat{z}$$

The above is really 3 functions associated with each of the 3 spatial coordinate directions. Under a rotation through an angle θ , about the \hat{z} axis the functions transform from, $f_i(x, y, z)$ to $f'_i(x', y', z')$. The transformation is written;

$$f'_i = \sum_j a_{ij} f_j$$

$$a_{ij} \rightarrow \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Generally one can rotate about any axis, and the transformation for other axes is an obvious permutation. As an aside here, note that a transformation about an arbitrary axis can be obtained by a rotation about the three Euler angles in succession α, β, γ as illustrated in Figure 1. Thus, there is a transformation a_{ij} which takes the function, \vec{F} , defined in the coordinate frame x, y, z into the function \vec{F}' in the coordinate frame x', y', z' .

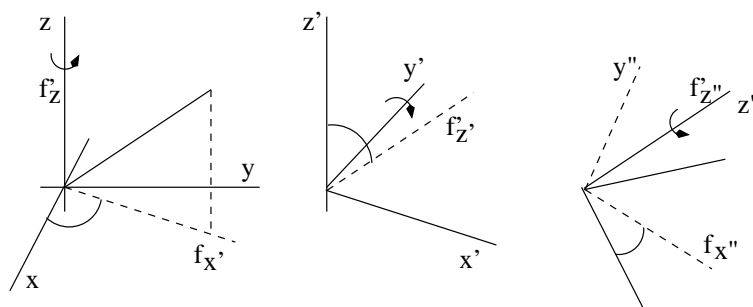


Figure 1: An example to obtain the rotation of a coordinate frame about an arbitrary axis defined by using the Euler angles

$$a_{ij} \rightarrow \begin{pmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix} \\ \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The rotation matrices are unitary. That is they preserve the length of the rotated vector. Multiply the matrix by its transpose to show this. Note that tensor operations reproduce matrix algebra as the product below indicates multiplication of a row vector by a column vector.

$$\vec{F} \cdot \vec{F}^\dagger = \sum_i \sum_j a_{ij} f_j \sum_k a_{ki}^\dagger f_k^\dagger \\ \vec{F} \cdot \vec{F}^\dagger = \sum_{jk} f_j f_k^\dagger \sum_i a_{ij} a_{ki}^\dagger \\ \sum_i a_{ij} a_{ki}^\dagger = \delta_{jk}$$

The matrix obtained in the above when summed over the direct product of the a_{ij} , is a diagonal matrix with 1's along the diagonal. This matrix has the determinant of +1. Now consider an operation which preserves length but reflects the vector about the origin. That is, the transformation which changes a right handed coordinate system into a left handed one.

$$x \rightarrow -x \quad y \rightarrow -y \quad z \rightarrow -z$$

This forms a matrix which is diagonal with -1's along the diagonal. It has a determinant of -1, and compared to the above matrix, represents the symmetry operation of parity. A vector changes sign upon reflection, *ie* a vector in a right handed frame is a negative vector in a left handed frame. If the vector does not change sign under reflection then it is a pseudo-vector or axial-vector. The cross product of two real vectors is a pseudo-vector as is obvious since each vector changes sign upon reflection and their product does not. The operations of rotation and reflection can be written as components of matrices, and as such, represent linear transformations of the vector components.

Remember that a scalar function remains constant under rotations and also reflections. Since a scalar is represented by one function, the transformation is a 1-D matrix or just a number, in this case 1. However, if the scalar function changes sign under reflection, it is a pseudo-scalar. Using the concepts of vector and scalar, linear transformations can be expanded to higher dimensions.

The fundamental characteristic of a tensor is the linear connection between the differential forms of a function in different coordinate frames.

Thus a tensor of rank 0 is a scalar (or pseudo-scalar) and a tensor of rank 1 is a vector (or pseudo-vector). A tensor of higher rank, $F_{k,l,\dots}$, transforms the function, $F_{k,l,\dots}$ as;

$$F'_{i,j,\dots} = \sum_{k,l,\dots} a_{k,l,\dots} F_{k,l,\dots}$$

2 Summation Convention

Suppose n independent variables, x^i with $i = 1, 2, \dots, n$. The set of values x^i define a point in a n -dimensional space. Note the superscript index. This does not mean the variable is to be raised to the i^{th} power. This notation will become obvious later. There are n independent functions, $\phi^i(x^1, x^2, \dots, n)$ in this space. For these functions to be linearly independent, the Jacobian cannot vanish and any other function in the space can be described by a linear combination of functions which “span” the space. Therefore;

$$J = \begin{vmatrix} \frac{\partial \phi^1}{\partial x^1} & \cdots & \frac{\partial \phi^n}{\partial x^1} \\ \frac{\partial \phi^1}{\partial x^n} & \cdots & \frac{\partial \phi^n}{\partial x^n} \end{vmatrix} \neq 0$$

Let $x'^i = \phi^i$ define another coordinate system, so in the same way which previously developed the scale factors;

$$\frac{\partial x^k}{\partial x^j} = \sum_i \frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^j} = \delta_j^k$$

A direction at a point is;

$$dx'^i = \sum_j \frac{\partial x'^i}{\partial x^j} dx^j$$

Now introduce the summation convention which assumes that a repeated index (a dummy variable) is summed. The summation sign will be dropped and summation indicated by a repeated index, so the partial derivative equations above are written as;

$$\frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^j} = \delta_j^k$$

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j$$

3 Contravariant and Covariant vectors

Now use both a superscript and a subscript on the variable to represent a tensor component. This is to denote different types of tensors. Note that a distance transforms as (summation convention);

$$dx'^i = \frac{\partial x'^i}{\partial x^j} dx^j$$

Thus this form represents a true vector, with components, say A^i , and it has the linear transform;

$$A'^i = \frac{\partial x'^i}{\partial x^j} A^j$$

To clarify, the above transformation describes how a component of a **contravariant** vector transforms.

Next consider the gradient (vector) operator;

$$\vec{\nabla}\phi = \frac{\partial\phi}{\partial x^i} \hat{x}^i$$

This form transforms as ;

$$\frac{\partial\phi'}{\partial x'^i} = \frac{\partial\phi}{\partial x^j} \frac{\partial x^j}{\partial x'^i}$$

Note that the gradient transforms as $\frac{\partial x^j}{\partial x'^i}$ rather than $\frac{\partial x'^i}{\partial x^j}$. This transforms a covariant vector component which will be noted by a subscript.

$$A'_i = \frac{\partial x^j}{\partial x'^i} A_j$$

In a transformation between Cartesian coordinate frames, there is no difference between contravariant and covariant forms.

$$\frac{\partial x^i}{\partial x'^j} = \frac{\partial x'^i}{\partial x^j} = a_j^i = a_i^j$$

In a general curvilinear coordinate frame, this is not the case. Therefore unless stated differently, a superscript is used to denote a contravariant component of a tensor and a subscript to denote a covariant component. A contravariant vector determines a direction and magnitude of a displacement at some point in space. It forms a vector field. A covariant vector describes the change in a particular direction of the field at a point in space.

Let λ^i be any n functions of the coordinates, x^j . The contravariant transformation is;

$$\lambda'^i = \frac{\partial x'^i}{\partial x^j} \lambda^j$$

Then write, since $\frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^j} = \delta_j^k$;

$$\frac{\partial x^k}{\partial x'^i} \lambda'^i = \lambda^j \frac{\partial x^k}{\partial x'^i} \frac{\partial x'^i}{\partial x^j} = \lambda^k$$

Thus we have found the inverse transformation for components of a **covariant vector**. Rewrite the above relations for vector components λ^j and μ_k and compose the form (summation convention);

$$\lambda'^j \mu'_j = \lambda^j \frac{\partial x'^i}{\partial x^j} \mu_k \frac{\partial x^k}{\partial x'^i} = \lambda^j \mu_k \delta_j^k = \lambda^j \mu_j$$

This is an invariant of the transformation. The result is a scalar which is invariant under coordinate transformations. In the old way of combining vectors this is just the dot or scalar product. Note that this is a summation over the product of contravariant and covariant components of two vectors.

4 Tensors

To begin, define two contravariant vectors, λ^i and η^i . Study the transformations from unprimed to a primed coordinate frame. Also define two covariant vectors, μ_i , and ζ_i . The following forms are to be used;

$$A^{ij} = \lambda^i \eta^j$$

$$A_{ij} = \mu_i \zeta_j$$

$$A_j^i = \lambda^i \mu_j$$

Then define combinations of the primed forms in a similar way. The transformation matrices between these are;

$$A'^{ij} = A^{kl} \frac{\partial x'^i}{\partial x^k} \frac{\partial x'^j}{\partial x^l}$$

$$A'_{ij} = A_{kl} \frac{\partial x^k}{\partial x'^i} \frac{\partial x^l}{\partial x'^j}$$

$$A_j'^i = A_l^k \frac{\partial x'^i}{\partial x^k} \frac{\partial x^l}{\partial x'^j}$$

These forms are tensors of 2^{nd} order. A^{kl} is a contravariant tensor, A_{kl} is a covariant tensor, and A_l^k is a mixed tensor. Note that there are n^2 elements in each tensor. The Kronecker delta, δ_j^k , is a mixed tensor of 2^{nd} order.

$$\delta_l^k = \frac{\partial x'^k}{\partial x^j} \frac{\partial x^j}{\partial x'^l}$$

Tensors of any order may be constructed in a similar way. The above construction uses the direct product of a tensor of lower order to produce one of higher order. Remember tensors are defined by their linear transformation properties. Thus;

$$A'^{i,j,\dots,k} = A^{m,n,\dots,p} \left(\frac{\partial x'^i}{\partial x^m} \right) \dots \left(\frac{\partial x'^k}{\partial x^p} \right)$$

The components of covariant and mixed tensors transform in a similar way. However, note that the order of the indicies (not the order of the terms in parenthesis) is important. Consider an arbitrary contravariant tensor A^{ij} . Write this as;

$$A^{ij} = (1/2)[A^{ij} + A^{ji}] + (1/2)[A^{ij} - A^{ji}]$$

Addition/subtraction of tensors can only be done for tensors of the same order. The first term on the right side in the above equation forms a symmetric tensor and the second an anti-symmetric tensor. Then an arbitrary tensor is a combination of symmetric and anti-symmetric tensors.

The concept of higher dimensional tensors is developed by directly multiplying two tensors of lower order to obtain a tensor of higher order. This is the outer (direct) product of the tensors. The inner (dot or scalar) product of two tensors forms a tensor of lower order. Thus consider the inner product of the tensors A_{ij} and B^{jkl} . The inner product sums over the repeated index, in this case, j , to get a tensor of rank 3.

$$A'_{ij} B'^{jkl} = \Gamma_i^{kl}$$

The direct product is

$$\Gamma_{uv}^{rkl} = A_{uv} B^{rkl}$$

Note that;

$$\left(\frac{\partial x^j}{\partial x'^w} \right) \left(\frac{\partial x'^w}{\partial x^t} \right) = \delta_t^j$$

This results in ;

$$A'_{ij} B'^{jkl} = \Gamma_r^{uv} \left(\frac{\partial x^r}{\partial x'^i} \right) \left(\frac{\partial x'^k}{\partial x'^u} \right) \left(\frac{\partial x'^l}{\partial x'^v} \right)$$

The above transforms as a mixed tensor of rank 3. The process of reducing the order of a tensor by the inner product is called contraction.

5 The metric

The metric is used in determining the differential length in a coordinate frame.

$$ds^2 = g_{ij} d\eta^i d\eta^j$$

$$g_{jk} = \sum_i \frac{\partial x^i}{\partial \eta_j} \frac{\partial x^i}{\partial \eta_k}$$

Perhaps a more insightful way to define the metric is to consider a length element, ds , along the i^{th} coordinate, x^i , which has a basis unit vector direction, \hat{e}_i .

$$ds^i = \hat{e}_i dx^i$$

Here ds^i is a length in the i^{th} direction. Now suppose a reciprocal unit vector basis, \hat{e}_i such that $\hat{e}^i \hat{e}_j = \delta_j^i$. Then contract the length vectors and assume the summation convention.

$$d\vec{s} \cdot d\vec{s} = ds^2 = dx_i dx_j \hat{e}^i \hat{e}^j = dx_i dx_j g^{ij}$$

In the above, g^{ij} is the contravariant metric matrix. It is symmetric and depends on the coordinate position. Note that ds^2 is a scalar length produced by a contraction of the two tensors. Now the reciprocal set of unit basis vectors can be expanded in terms of the unit basis vectors of the space.

$$\hat{e}_i = c_{ij} \hat{e}^j \quad (\text{Summation convention})$$

Then $\hat{e}_j \hat{e}_i = c_{ij} = g_{ij}$, and $e_i = g_{ij} \hat{e}^j$ and $e^i = g^{ij} e_j$. The metric is used to raise or lower an index in a tensor. Also, the metric is restricted so that $|g_{ij}| \neq 0$, but otherwise ds^2 can be < 0 so we must take $|ds^2|$ to be the measure of length. The metric is still defined as;

$$ds^2 = g_{ij} dx^i dx^j$$

An arc of length between a and b is;

$$s = \int_a^b dt \sqrt{\alpha g_{ij} \frac{\partial x^i}{\partial t} \frac{\partial x^j}{\partial t}}$$

The factor, α , is taken as ± 1 so that each element in the sum is > 0 , and t is the parametric variable.

6 Conjugate Tensors

The metric g_{ij} is a symmetric covariant tensor of order 2. Thus $g_{ij} = g_{ji}$. Further we have that;

$$g = \begin{vmatrix} g_{11} & \cdots & g_{1n} \\ & \cdots & \\ g_{n1} & \cdots & g_{nn} \end{vmatrix}$$

Then let g^{ij} be the cofactor of the element g_{ij} divided by g . The cofactor of a tensor element, A_{ij} , is given by;

$$Cofactor(A_{ij}) = (-1)^{i+j} M_{ij}$$

where M_{ij} is the minor of the element A_{ij} . The minor is obtained from the determinant of the tensor after deleting the row and column containing the element. This means that the determinant of the tensor A is;

$$|A| = \sum_{k=1}^n A^{ij} (-1)^{i+j} M_{ij}$$

This then gives the relation;

$$g^{ij} g_{kj} = \delta_k^i$$

The g^{ij} are elements of a contravariant tensor of order 2. This tensor is the conjugate of g_{ij} . Using this symmetric tensor, one may obtain a tensor of the same order but of different character (raise or lower the index - ie change from covariant to contravariant or contravariant to covariant).

$$A_{jk}^l = g^{li} A_{ijk}$$

$$A^{lmp} = g^{li} g^{mj} g^{pk} A_{ijk}$$

Note that the process is reversible. Also;

$$\sum_j g^{ij} g_{kj} = (1/g) \sum_j (-1)^{i+j} M_{ij} g_{kj}$$

Unless $i = k$ the product of terms is obtained by using one term of one row and cofactors of

terms in another row. In the case of $i = k$ the determinant results. As a result;

$$a_i = g_{ij}a^j$$

$$a^j = g^{ij}a_i$$

7 Levi-Civita symbol

Define the following tensor of rank 3.

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{321} = 1$$

$$\epsilon_{132} = \epsilon_{213} = \epsilon_{312} = -1$$

$$\text{All other } \epsilon = 0$$

This is a 3-D Levi-Civita tensor and is obviously a pseudo-tensor as odd permutations of the indicies are -1 times the even permutations. We can also define the conjugate tensor ϵ^{ijk} . Suppose a set of vectors $\lambda^i \eta^j \zeta^i$ with $i = 1, 2, 3$. Contract a tensor with the Levi-Civita tensor to produce the pseudo-scalar, ϕ .

$$\phi = \epsilon_{ijk} \lambda^i \eta^j \zeta^k$$

$$\phi = \lambda^1 \eta^2 \zeta^3 - \lambda^1 \eta^3 \zeta^2 + \lambda^2 \eta^3 \zeta^1 - \lambda^2 \eta^1 \zeta^3 + \lambda^3 \eta^1 \zeta^2 - \lambda^3 \eta^2 \zeta^1$$

The Levi-Civita tensor may be expanded to rank 4, ϵ_{ijkl} , by a similar definition. If we allow the vectors $\lambda^i \eta^j \zeta^i$ to be differential vectors lying along each of the Cartesian coordinate axes, *ie* $\lambda \rightarrow (dx, 0, 0)$, $\eta \rightarrow (0, dy, 0)$, $\zeta \rightarrow (0, 0, dz)$ The value of $\phi = dx dy dz = d\tau$. This is a differential volume element, which we identify as a pseudo-scalar. A pseudotensor of odd rank does not change sign under parity inversion while one of even rank does.

8 Dual tensors

Any anti-symmetric tensor of rank greater than 2 can have a dual representation. This is understood by contracting the tensor with a Levi-Civita tensor. Thus suppose a tensor of rank 3. To be specific, assume the angular momentum obtained by;

$$L^{ij} \rightarrow (1/2) \begin{pmatrix} 0 & x^1 p^2 - x^2 p^1 & -x^1 p^3 + p^3 x^1 \\ -x^1 p^2 + x^2 p^1 & 0 & x^2 p^3 - p^2 x^3 \\ x^1 p^3 - x^3 p^1 & -x^2 p^3 + p^3 x^2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & C_{12} & -C_{31} \\ -C_{21} & 0 & C_{23} \\ C_{31} & -C_{32} & 0 \end{pmatrix}$$

where to be anti-symmetric $C_{ij} = -C_{ji}$. Then the contraction with the 3-D Levi-Civita tensor yields a vector.

$$c'_i = (1/2)\epsilon_{ijk}C_{jk}$$

This produces, for the angular momentum example;

$$\vec{L} = [yp_z - zp_y]\hat{x} - [xp_z - zp_x]\hat{y} + [yp_x - xp_y]\hat{z}$$

Note \vec{C}' transforms as a vector as previously proved by tensor contraction, but this vector is a pseudo-vector because of its symmetry properties. Note that the cross product is a special anti-symmetric tensor with a dual which has the rotation properties of a vector, but the symmetry under inversion is a tensor of rank 2. Tensor properties may also be developed from the concept of the increasing order of surfaces. The properties of a point are equivalent to a scalar, the properties of a line are vectors, the properties of a surface are pseudo-vectors, etc.

9 Covariant derivative

The covariant derivative assures that a vector is independent of its description in an arbitrary coordinate frame. That is, the vector has a magnitude and points in a specific direction independent of the frame of reference. The covariant derivative removes the change in the vector due the curvature of the coordinate frame. Therefore the components of a covariant vector representing the rate of change of an ordinary vector, \vec{A} , with respect to the η^i axis are;

$$A^j_{,i} = \frac{\partial A^j}{\partial \eta_i} + \sum_k A^k \Gamma^j_{ki}$$

The symbol Γ^j_{ki} is a Christoffel symbol, and represents the change in the direction of the unit vectors. It will be discussed in more detail, but for the moment represents the projection of the change in the unit vectors of the space as a vector is displaced a distance, dq^j .

$$d\hat{e}_i = \Gamma^k_{ij}\hat{e}_k dq^j$$

The comma indicates that $A^j_{,i}$ is the covariant derivative. The $A^i_{,j}$ are the components of a mixed tensor, covariant with respect to the index j and contravariant with respect to

the index i . Covariant differentiation can also be extended to tensors of higher order.

The contracted tensor, $\sum_n A^n_{,n}$, represents the *Div* of the vector \vec{A} ;

$$\sum_n A^n_{,n} = \sum_n \frac{\partial A^n}{\partial \eta_n} + \sum_{n,m} A^n \Gamma_{nm}^m = \sum_n \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial \eta_n} (A^n h_1 h_2 h_3)$$

Identify $A_n = F^n/h_n$. Substitution in the above yields *Div* \vec{A} .

$$Div \vec{A} = \frac{1}{h_1 h_2 h_3} \sum_n \frac{\partial}{\partial \eta_n} \left[\frac{h_1 h_2 h_3 A^n}{h_n} \right]$$

The curl of vector \vec{B} , can also be developed for orthogonal coordinates by consideration of the component $A^i = -\left(\frac{1}{h_1 h_2 h_3}\right)[B_{j,k} - B_{k,j}]$. The scalar Laplacian can also be written as a covariant derivative, $\sum_n \left(\frac{1}{h_n^2} \frac{\partial \phi}{\partial \eta_n}\right)_{,n}$. Thus the use of the covariant derivative allows one to express an equation in the same form for any coordinate system.

10 Geodesic lines

In a general curvilinear coordinate system (Riemann space) there is a unique shortest line which connects two points. As you know, in a spherical system this line is a great circle. These lines are called geodesics. The length is given by;

$$s = \int_a^b ds = \int_a^b dt [g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}]^{1/2}$$

In the above t is the parametric variable. To find the minimum length, apply the calculus of variations. This will be discussed later in the course. Basically one varies all paths between a and b which go through the end points a and b , choosing the one which is stationary, *ie* this is similar to using the derivative of a function to determine an extremum. The operation reduces to the Euler-Lagrange equations and finally the differential equation;

$$\frac{d^2 x^l}{ds^2} + \Gamma_{ik}^l \frac{dx^i}{ds} \frac{dx^k}{ds} = 0$$

In a Cartesian system the Christoffel symbol vanishes which leads to a linear relation between the length and the coordinates.

Although this will be discussed in much more detail in future lectures, an example is the equations of general relativity. We live in a 4-dimensional Riemann (curvilinear) space as shown in Figure 2. This figure illustrates that the world line is a geodesic in 4-D space. In

the figure, τ , is the proper time. A line element is written, $ds = \hat{a}_i d\eta^i$, so that the length squared is;

$$ds^2 = g_{ij}d\eta_i d\eta^j$$

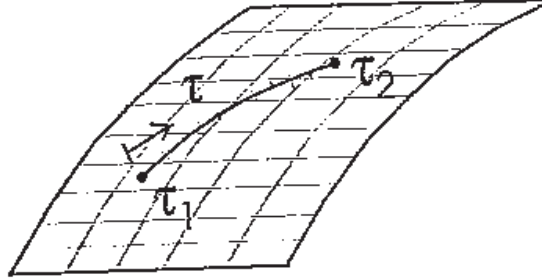


Figure 2: A 4-D Riemann space showing a world line with proper time, τ along the path

In a free, flat space, ie a vacuum where there is no matter to introduce curvature, the metric is given by;

$$g = \begin{vmatrix} g_{11} & \cdots & g_{1n} \\ & \cdots & \\ g_{n1} & \cdots & g_{nn} \end{vmatrix}$$

At each point in the space, the structure is given by the Einstein tensor equation;

$$G^{uv} = \kappa T^{uv}$$

In the above, G^{uv} is derived from the curvature tensor of space, and T^{uv} is the energy-momentum tensor. The path of a world line in this space is given by a geodesic. The curvature tensor is written in more detail as;

$$G^{uv} = R^{uv} - 1/2 Rg^{uv} \quad R = R^u_u \quad R_{uv} = R^w_{uvw}$$

$$R^w_{upv} = \left[\frac{\partial \Gamma^w_{uv}}{\partial \eta^p} + \Gamma^w_{ps} \Gamma^s_{uv} \right] - \left[\frac{\partial \Gamma^w_{uv}}{\partial \eta^v} + \Gamma^w_{vs} \Gamma^s_{up} \right]$$

with the affine connection representing the differential change in the basis vectors with respect to a variation of the coordinates;

$$\Gamma^w_{uv} = (1/2)g^{ws} \left[\frac{\partial g_{sv}}{\partial \eta^u} + \frac{\partial g_{su}}{\partial \eta^v} - \frac{\partial g_{uv}}{\partial \eta^s} \right]$$

The solution for a flat space is easily obtained as the Christoffel symbols vanish. One can also look for perturbative solutions where the flat metric is slightly perturbed. An analytic solution of the general form is however, not possible.