GENERAL THEORY OF RELATIVITY

P. A. M. DIRAC
Florida State University

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS, New York · London · Sydney · Toronto

Library of Congress Cataloging in Publication Data

Dirac, Paul Adrien Maurice, 1902–
General theory of relativity.

"A Wiley-Interscience publication."
Based on a course of lectures given at Florida State University, Physics Dept.
Includes index.

1. General relativity (Physics) I. Title.

QC173.6.D57 530.1'1 75–8690
ISBN 0-471-21575-9

Printed in the United States of America

10 9 8 7 6 5 4

Preface

Einstein's general theory of relativity requires a curved space for the description of the physical world. If one wishes to go beyond a superficial discussion of the physical relations one needs to set up precise equations for handling curved space. There is a well-established but rather complicated mathematical technique that does this. It has to be mastered by any student who wishes to understand Einstein's theory.

This book is built up from a course of lectures given at the Physics Department of Florida State University and has the aim of presenting the indispensable material in a direct and concise form. It does not require previous knowledge beyond the basic ideas of special relativity and the handling of differentiations of field functions. It will enable the student to pass through the main obstacles of understanding general relativity with the minimum expenditure of time and trouble and to become qualified to continue more deeply into any specialized aspects of the subject that interest him.

P. A. M. DIRAC

Tallahassee, Florida
February 1975

Contents

1. Special relativity

For the space-time of physics we need four coordinates, the time t and three space coordinates x, y, z. We put

$$t = x^0, \qquad x = x^1, \qquad y = x^2, \qquad z = x^3,$$

so that the four coordinates may be written x^μ, where the suffix μ takes on the four values 0, 1, 2, 3. The suffix is written in the upper position in order that we may maintain a "balancing" of the suffixes in all the general equations of the theory. The precise meaning of balancing will become clear a little later.

Let us take a point close to the point that we originally considered and let its coordinates be $x^\mu + dx^\mu$. The four quantities dx^μ which form the displacement may be considered as the components of a vector. The laws of special relativity allow us to make linear nonhomogeneous transformations of the coordinates, resulting in linear homogeneous transformations of the dx^μ. These are such that, if we choose units of distance and of time such that the velocity of light is unity,

$$(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \tag{1.1}$$

is invariant.

Any set of four quantities A^μ that transform under a change of coordinates in the same way as the dx^μ form what is called a *contravariant vector*. The invariant quantity

$$(A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 = (A, A) \tag{1.2}$$

may be called the squared length of the vector. With a second contravariant vector B^μ, we have the scalar product invariant

$$A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3 = (A, B). \tag{1.3}$$

In order to get a convenient way of writing such invariants we introduce the device of lowering suffixes. Define

$$A_0 = A^0, \qquad A_1 = -A^1, \qquad A_2 = -A^2, \qquad A_3 = -A^3. \tag{1.4}$$

Then the expression on the left-hand side of (1.2) may be written $A_\mu A^\mu$, in which it is understood that a summation is to be taken over the four values of μ. With the same notation we can write (1.3) as $A_\mu B^\mu$ or else $A^\mu B_\mu$.

The four quantities A_μ introduced by (1.4) may also be considered as the components of a vector. Their transformation laws under a change of coordinates are somewhat different from those of the A^μ, because of the differences in sign, and the vector is called a *covariant vector.*

From the two contravariant vectors A^μ and B^μ we may form the sixteen quantities $A^\mu B^\nu$. The suffix ν, like all the Greek suffixes appearing in this work, also takes on the four values 0, 1, 2, 3. These sixteen quantities form the components of a tensor of the second rank. It is sometimes called the outer product of the vectors A^μ and B^μ, as distinct from the scalar product (1.3), which is called the inner product.

The tensor $A^\mu B^\nu$ is a rather special tensor because there are special relations between its components. But we can add together several tensors constructed in this way to get a general tensor of the second rank; say

$$T^{\mu\nu} = A^\mu B^\nu + A'^\mu B'^\nu + A''^\mu B''^\nu + \cdots. \tag{1.5}$$

The important thing about the general tensor is that under a transformation of coordinates its components transform in the same way as the quantities $A^\mu B^\nu$.

We may lower one of the suffixes in $T^{\mu\nu}$ by applying the lowering process to each of the terms on the right-hand side of (1.5). Thus we may form $T_\mu{}^\nu$ or $T^\mu{}_\nu$. We may lower both suffixes to get $T_{\mu\nu}$.

In $T_\mu{}^\nu$ we may set $\nu = \mu$ and get $T_\mu{}^\mu$. This is to be summed over the four values of μ. A summation is always implied over a suffix that occurs twice in a term. Thus $T_\mu{}^\mu$ is a scalar. It is equal to $T^\mu{}_\mu$.

We may continue this process and multiply more than two vectors together, taking care that their suffixes are all different. In this way we can construct tensors of higher rank. If the vectors are all contravariant, we get a tensor with all its suffixes upstairs. We may then lower any of the suffixes and so get a general tensor with any number of suffixes upstairs and any number downstairs.

We may set a downstairs suffix equal to an upstairs one. We then have to sum over all values of this suffix. The suffix becomes a dummy. We are left with a tensor having two fewer effective suffixes than the original one. This process is called *contraction*. Thus, if we start with the fourth-rank tensor $T^\mu{}_{\nu\rho}{}^\sigma$, one way of contracting it is to put $\sigma = \rho$, which gives the second rank tensor $T^\mu{}_{\nu\rho}{}^\rho$, having only sixteen components, arising from the four values of μ and ν. We could contract again to get the scalar $T^\mu{}_{\mu\rho}{}^\rho$, with just one component.

At this stage one can appreciate the balancing of suffixes. Any effective suffix occurring in an equation appears once and only once in each term of the equation, and always upstairs or always downstairs. A suffix occurring twice in a term is a dummy, and it must occur once upstairs and once downstairs. It may be replaced by any other Greek letter not already mentioned in the term. Thus $T^{\mu}{}_{\nu\rho}{}^{\rho} = T^{\mu}{}_{\nu\alpha}{}^{\alpha}$. A suffix must never occur more than twice in a term.

2. Oblique axes

Before passing to the formalism of general relativity it is convenient to consider an intermediate formalism—special relativity referred to oblique rectilinear axes.

If we make a transformation to oblique axes, each of the dx^{μ} mentioned in (1.1) becomes a linear function of the new dx^{μ} and the quadratic form (1.1) becomes a general quadratic form in the new dx^{μ}. We may write it

$$g_{\mu\nu}\, dx^{\mu}\, dx^{\nu}, \tag{2.1}$$

with summations understood over both μ and ν. The coefficients $g_{\mu\nu}$ appearing here depend on the system of oblique axes. Of course we take $g_{\mu\nu} = g_{\nu\mu}$, because any difference of $g_{\mu\nu}$ and $g_{\nu\mu}$ would not show up in the quadratic form (2.1). There are thus ten independent coefficients $g_{\mu\nu}$.

A general contravariant vector has four components A^{μ} which transform like the dx^{μ} under any transformation of the oblique axes. Thus

$$g_{\mu\nu} A^{\mu} A^{\nu}$$

is invariant. It is the squared length of the vector A^{μ}.

Let B^{μ} be a second contravariant vector; then $A^{\mu} + \lambda B^{\mu}$ is still another, for any value of the number λ. Its squared length is

$$g_{\mu\nu}(A^{\mu} + \lambda B^{\mu})(A^{\nu} + \lambda B^{\nu}) = g_{\mu\nu} A^{\mu} A^{\nu} + \lambda(g_{\mu\nu} A^{\mu} B^{\nu} + g_{\mu\nu} A^{\nu} B^{\mu}) + \lambda^2 g_{\mu\nu} B^{\mu} B^{\nu}.$$

This must be an invariant for all values of λ. It follows that the term independent of λ and the coefficients of λ and λ^2 must separately be invariants. The coefficient of λ is

$$g_{\mu\nu}A^{\mu}B^{\nu} + g_{\mu\nu}A^{\nu}B^{\mu} = 2g_{\mu\nu}A^{\mu}B^{\nu},$$

since in the second term on the left we may interchange μ and ν and then use $g_{\mu\nu} = g_{\nu\mu}$. Thus we find that $g_{\mu\nu}A^{\mu}B^{\nu}$ is an invariant. It is the scalar product of A^{μ} and B^{μ}.

Let g be determinant of the $g_{\mu\nu}$. It must not vanish; otherwise the four axes would not provide independent directions in space-time and would not be suitable as axes. For the orthogonal axes of the preceding section the diagonal elements of $g_{\mu\nu}$ are $1, -1, -1, -1$ and the nondiagonal elements are zero. Thus $g = -1$. With oblique axes g must still be negative, because the oblique axes can be obtained from the orthogonal ones by a continuous process, resulting in g varying continuously, and g cannot pass through the value zero.

Define the covariant vector A_{μ}, with a downstairs suffix, by

$$A_{\mu} = g_{\mu\nu}A^{\nu}. \tag{2.2}$$

Since the determinant g does not vanish, these equations can be solved for A^{ν} in terms of the A_{μ}. Let the result be

$$A^{\nu} = g^{\mu\nu}A_{\mu}. \tag{2.3}$$

Each $g^{\mu\nu}$ equals the cofactor of the corresponding $g_{\mu\nu}$ in the determinant of the $g_{\mu\nu}$, divided by the determinant itself. It follows that $g^{\mu\nu} = g^{\nu\mu}$.

Let us substitute for the A^{ν} in (2.2) their values given by (2.3). We must replace the dummy μ in (2.3) by some other Greek letter, say ρ, in order not to have three μ's in the same term. We get

$$A_{\mu} = g_{\mu\nu}g^{\nu\rho}A_{\rho}.$$

Since this equation must hold for any four quantities A_{μ}, we can infer

$$g_{\mu\nu}g^{\nu\rho} = g_{\mu}^{\rho}, \tag{2.4}$$

where

$$\begin{aligned} g_{\mu}^{\rho} &= 1 \quad \text{for } \mu = \rho, \\ &= 0 \quad \text{for } \mu \neq \rho. \end{aligned} \tag{2.5}$$

The formula (2.2) may be used to lower any upper suffix occurring in a tensor. Similarly, (2.3) can be used to raise any downstairs suffix. If a suffix is lowered and raised again, the result is the same as the original tensor, on account of (2.4) and (2.5). Note that g^ρ_μ just produces a substitution of ρ for μ,

$$g^\rho_\mu A^\mu = A^\rho,$$

or of μ for ρ,

$$g^\rho_\mu A_\rho = A_\mu.$$

If we apply the rule for raising a suffix to the μ in $g_{\mu\nu}$, we get

$$g^\alpha{}_\nu = g^{\alpha\mu} g_{\mu\nu}.$$

This agrees with (2.4), if we take into account that in $g^\alpha{}_\nu$ we may write the suffixes one above the other because of the symmetry of $g_{\mu\nu}$. Further we may raise the suffix ν by the same rule and get

$$g^{\alpha\beta} = g^{\nu\beta} g^\alpha_\nu,$$

a result which follows immediately from (2.5). The rules for raising and lowering suffixes apply to all the suffixes in $g_{\mu\nu}$, g^μ_ν, $g^{\mu\nu}$.

3. Curvilinear coordinates

We now pass on to a system of curvilinear coordinates. We shall deal with quantities which are located at a point in space. Such a quantity may have various components, which are then referred to the axes at that point. There may be a quantity of the same nature at all points of space. It then becomes a field quantity.

If we take such a quantity Q (or one of its components if it has several), we can differentiate it with respect to any of the four coordinates. We write the result

$$\frac{\partial Q}{\partial x^\mu} = Q_{,\mu}.$$

A downstairs suffix preceded by a comma will always denote a derivative in this way. We put the suffix μ downstairs in order to balance the upstairs μ

in the denominator on the left. We can see that the suffixes balance by noting that the change in Q, when we pass from the point x^μ to the neighboring point $x^\mu + \delta x^\mu$, is

$$\delta Q = Q_{,\mu}\, \delta x^\mu. \tag{3.1}$$

We shall have vectors and tensors located at a point, with various components referring to the axes at that point. When we change our system of coordinates, the components will change according to the same laws as in the preceding section, depending on the change of axes at the point concerned. We shall have a $g_{\mu\nu}$ and a $g^{\mu\nu}$ to lower and raise suffixes, as before. But *they are no longer constants*. They vary from point to point. They are field quantities.

Let us see the effect of a particular change in the coordinate system. Take new curvilinear coordinates x'^μ, each a function of the four x's. They may be written more conveniently $x^{\mu'}$, with the prime attached to the suffix rather than the main symbol.

Making a small variation in the x^μ, we get the four quantities δx^μ forming the components of a contravariant vector. Referred to the new axes, this vector has the components

$$\delta x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\nu}\, \delta x^\nu = x^{\mu'}_{,\nu}\, \delta x^\nu,$$

with the notation of (3.1). This gives the law for the transformation of any contravariant vector A^ν; namely,

$$A^{\mu'} = x^{\mu'}_{,\nu} A^\nu. \tag{3.2}$$

Interchanging the two systems of axes and changing the suffixes, we get

$$A^\lambda = x^\lambda_{,\mu'} A^{\mu'}. \tag{3.3}$$

We know from the laws of partial differentiation that

$$\frac{\partial x^\lambda}{\partial x^{\mu'}} \frac{\partial x^{\mu'}}{\partial x^\nu} = g^\lambda_\nu,$$

with the notation (2.5). Thus

$$x^\lambda_{,\mu'} x^{\mu'}_{,\nu} = g^\lambda_\nu. \tag{3.4}$$

This enables us to see that the two equations (3.2) and (3.3) are consistent, since if we substitute (3.2) into the right-hand side of (3.3), we get

$$x^\lambda_{,\mu'} x^{\mu'}_{,\nu} A^\nu = g^\lambda_\nu A^\nu = A^\lambda.$$

To see how a covariant vector B_μ transforms, we use the condition that $A^\mu B_\mu$ is invariant. Thus with the help of (3.3)

$$A^{\mu'} B_{\mu'} = A^\lambda B_\lambda = x^\lambda_{,\mu'} A^{\mu'} B_\lambda.$$

This result must hold for all values of the four $A^{\mu'}$; therefore we can equate the coefficients of $A^{\mu'}$ and get

$$B_{\mu'} = x^\lambda_{,\mu'} B_\lambda. \tag{3.5}$$

We can now use the formulas (3.2) and (3.5) to transform any tensor with any upstairs and downstairs suffixes. We just have to use coefficients like $x^{\mu'}_{,\nu}$ for each upstairs suffix and like $x^\lambda_{,\mu'}$ for each downstairs suffix and make all the suffixes balance. For example

$$T^{\alpha'\beta'}{}_{\gamma'} = x^{\alpha'}_{,\lambda} x^{\beta'}_{,\mu} x^{\nu}_{,\gamma'} T^{\lambda\mu}{}_{\nu}. \tag{3.6}$$

Any quantity that transforms according to this law is a tensor. This may be taken as the definition of a tensor.

It should be noted that it has a meaning for a tensor to be symmetrical or antisymmetrical between two suffixes like λ and μ, because this property of symmetry is preserved with the change of coordinates.

The formula (3.4) may be written

$$x^\lambda_{,\alpha'} x^{\beta'}_{,\nu} g^{\alpha'}_{\beta'} = g^\lambda_\nu.$$

It just shows that g^λ_ν is a tensor. We have also, for any vectors A^μ, B^ν,

$$g_{\alpha'\beta'} A^{\alpha'} B^{\beta'} = g_{\mu\nu} A^\mu B^\nu = g_{\mu\nu} x^\mu_{,\alpha'} x^\nu_{,\beta'} A^{\alpha'} B^{\beta'}.$$

Since this holds for all values of $A^{\alpha'}$, $B^{\beta'}$, we can infer

$$g_{\alpha'\beta'} = g_{\mu\nu} x^\mu_{,\alpha'} x^\nu_{,\beta'}. \tag{3.7}$$

This shows that $g_{\mu\nu}$ is a tensor. Similarly, $g^{\mu\nu}$ is a tensor. They are called the *fundamental tensors.*

If S is any scalar field quantity, it can be considered either as a function of the four x^μ or of the four $x^{\mu'}$. From the laws of partial differentiation

$$S_{,\mu'} = S_{,\lambda} x^\lambda_{,\mu'}.$$

Hence the $S_{,\lambda}$ transform like the B_λ of equation (3.5) and thus *the derivative of a scalar field is a covariant vector field.*

4. Nontensors

We can have a quantity $N^{\mu}{}_{\nu\rho\ldots}$ with various up and down suffixes, which is not a tensor. If it is a tensor, it must transform under a change of coordinate system according to the law exemplified by (3.6). With any other law it is a nontensor. A tensor has the property that if all the components vanish in one system of coordinates, they vanish in every system of coordinates. This may not hold for a nontensor.

For a nontensor we can raise and lower suffixes by the same rules as for a tensor. Thus, for example,

$$g^{\alpha\nu}N^{\mu}{}_{\nu\rho} = N^{\mu\alpha}{}_{\rho}.$$

The consistency of these rules is quite independent of the transformation laws to a different system of coordinates. Similarly, we can contract a nontensor by putting an upper and lower suffix equal.

We may have tensors and nontensors appearing together in the same equation. The rules for balancing suffixes apply equally to tensors and nontensors.

THE QUOTIENT THEOREM

Suppose $P_{\lambda\mu\nu}$ is such that $A^{\lambda}P_{\lambda\mu\nu}$ is a tensor for any vector A^{λ}. Then $P_{\lambda\mu\nu}$ is a tensor.

To prove it, write $A^{\lambda}P_{\lambda\mu\nu} = Q_{\mu\nu}$. We are given that it is a tensor; therefore

$$Q_{\beta\gamma} = Q_{\mu'\nu'}x^{\mu'}_{,\beta}x^{\nu'}_{,\gamma}.$$

Thus

$$A^{\alpha}P_{\alpha\beta\gamma} = A^{\lambda'}P_{\lambda'\mu'\gamma'}x^{\mu'}_{,\beta}x^{\nu'}_{,\gamma}.$$

Since A^{λ} is a vector, we have from (3.2),

$$A^{\lambda'} = A^{\alpha}x^{\lambda'}_{,\alpha}.$$

So

$$A^{\alpha}P_{\alpha\beta\gamma} = A^{\alpha}x^{\lambda'}_{,\alpha}P_{\lambda'\mu'\nu'}x^{\mu'}_{,\beta}x^{\nu'}_{,\gamma}.$$

This equation must hold for all values of A^α, so

$$P_{\alpha\beta\gamma} = P_{\lambda'\mu'\nu'} x^{\lambda'}_{,\alpha} x^{\mu'}_{,\beta} x^{\nu'}_{,\gamma},$$

showing that $P_{\alpha\beta\gamma}$ is a tensor.

The theorem also holds if $P_{\lambda\mu\nu}$ is replaced by a quantity with any number of suffixes, and if some of the suffixes are upstairs.

5. Curved space

One can easily imagine a curved two-dimensional space as a surface immersed in Euclidean three-dimensional space. In the same way, one can have a curved four-dimensional space immersed in a flat space of a larger number of dimensions. Such a curved space is called a Riemann space. A small region of it is approximately flat.

Einstein assumed that physical space is of this nature and thereby laid the foundation for his theory of gravitation.

For dealing with curved space one cannot introduce a rectilinear system of axes. One has to use curvilinear coordinates, such as those dealt with in Section 3. The whole formalism of that section can be applied to curved space, because all the equations are local ones which are not disturbed by the curvature.

The invariant distance ds between a point x^μ and a neighboring point $x^\mu + dx^\mu$ is given by

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$$

like (2.1). ds is real for a timelike interval and imaginary for a spacelike interval.

With a network of curvilinear coordinates the $g_{\mu\nu}$, given as functions of the coordinates, fix all the elements of distance; so they fix the metric. They determine both the coordinate system and the curvature of the space.

6. Parallel displacement

Suppose we have a vector A^μ located at a point P. If the space is curved, we cannot give a meaning to a parallel vector at a different point Q, as one can easily see if one thinks of the example of a curved two-dimensional space in a three-dimensional Euclidean space. However, if we take a point P' close to P, there is a parallel vector at P', with an uncertainty of the second order, counting the distance from P to P' as the first order. Thus we can give a meaning to displacing the vector A^μ from P to P' keeping it parallel to itself and keeping the length constant.

We can transfer the vector continuously along a path by this process of parallel displacement. Taking a path from P to Q, we end up with a vector at Q which is parallel to the original vector at P with respect to this path. But a different path would give a different result. There is no absolute meaning to a parallel vector at Q. If we transport the vector at P by parallel displacement around a closed loop, we shall end up with a vector at P which is usually in a different direction.

We can get equations for the parallel displacement of a vector by supposing our four-dimensional physical space to be immersed in a flat space of a higher number of dimensions; say N. In this N-dimensional space we introduce rectilinear coordinates $z^n (n = 1, 2, \ldots, N)$. These coordinates do not need to be orthogonal, only rectilinear. Between two neighboring points there is an invariant distance ds given by

$$ds^2 = h_{nm}\, dz^n\, dz^m, \tag{6.1}$$

summed for $n, m = 1, 2, \ldots, N$. The h_{nm} are constants, unlike the $g_{\mu\nu}$. We may use them to lower suffixes in the N-dimensional space; thus

$$dz_n = h_{nm}\, dz^m.$$

Physical space forms a four-dimensional "surface" in the flat N-dimensional space. Each point x^μ in the surface determines a definite point y^n in the N-dimensional space. Each coordinate y^n is a function of the four x's; say $y^n(x)$. The equations of the surface would be given by eliminating the four x's from the $N y^n(x)$'s. There are $N - 4$ such equations.

By differentiating the $y^n(x)$ with respect to the parameters x^μ, we get

$$\frac{\partial y^n(x)}{\partial x^\mu} = y^n_{,\mu}.$$

For two neighboring points in the surface differing by δx^μ, we have

$$\delta y^n = y^n_{,\mu}\,\delta x^\mu. \tag{6.2}$$

The squared distance between them is, from (6.1)

$$\delta s^2 = h_{nm}\,\delta y^n\,\delta y^m = h_{nm}\,y^n_{,\mu}\,y^m_{,\nu}\,\delta x^\mu\,\delta x^\nu.$$

We may write it

$$\delta s^2 = y^n_{,\mu}\,y_{n,\nu}\,\delta x^\mu\,\delta x^\nu$$

on account of the h_{nm} being constants. We also have

$$\delta s^2 = g_{\mu\nu}\,\delta x^\mu\,\delta x^\nu.$$

Hence

$$g_{\mu\nu} = y^n_{,\mu}\,y_{n,\nu}. \tag{6.3}$$

Take a contravariant vector A^μ in physical space, located at the point x. Its components A^μ are like the δx^μ of (6.2). They will provide a contravariant vector A^n in the N-dimensional space, like the δy^n of (6.2). Thus

$$A^n = y^n_{,\mu}\,A^\mu. \tag{6.4}$$

This vector A^n, of course, lies in the surface.

Now shift the vector A^n, keeping it parallel to itself (which means, of course, keeping the components constant), to a neighboring point $x + dx$ in the surface. It will no longer lie in the surface at the new point, on account of the curvature of the surface. But we can project it on to the surface, to get a definite vector lying in the surface.

The projection process consists in splitting the vector into two parts, a tangential part and a normal part, and discarding the normal part. Thus

$$A^n = A^n_{\text{tan}} + A^n_{\text{nor}}. \tag{6.5}$$

Now if K^μ denotes the components of A^n_{tan} referred to the x coordinate system in the surface, we have, corresponding to (6.4),

$$A^n_{\text{tan}} = K^\mu y^n_{,\mu}(x + dx), \tag{6.6}$$

with the coefficients $y^n_{,\mu}$ taken at the new point $x + dx$.

A^n_{nor} is defined to be orthogonal to every tangential vector at the point $x + dx$, and thus to every vector like the right-hand side of (6.6), no matter what the K^μ are. Thus

$$A^n_{\text{nor}}\,y_{n,\mu}(x + dx) = 0.$$

If we now multiply (6.5) by $y_{n,\nu}(x + dx)$, the A^n_{nor} term drops out and we are left with

$$A^n y_{n,\nu}(x + dx) = K^\mu y^n_{,\mu}(x + dx) y_{n,\nu}(x + dx)$$
$$= K^\mu g_{\mu\nu}(x + dx)$$

from (6.3). Thus to the first order in dx

$$K_\nu(x + dx) = A^n[y_{n,\nu}(x) + y_{n,\nu,\sigma}\, dx^\sigma]$$
$$= A^\mu y^n_{,\mu}[y_{n,\nu} + y_{n,\nu,\sigma}\, dx^\sigma]$$
$$= A_\nu + A^\mu y^n_{,\mu} y_{n,\nu,\sigma}\, dx^\sigma.$$

This K_ν is the result of parallel displacement of A_ν to the point $x + dx$. We may put

$$K_\nu - A_\nu = dA_\nu,$$

so dA_ν denotes the change in A_ν under parallel displacement. Then we have

$$dA_\nu = A^\mu y^n_{,\mu} y_{n,\nu,\sigma}\, dx^\sigma. \tag{6.7}$$

7. Christoffel symbols

By differentiating (6.3) we get (omitting the second comma with two differentiations)

$$g_{\mu\nu,\sigma} = y^n_{,\mu\sigma} y_{n,\nu} + y^n_{,\mu} y_{n,\nu\sigma}$$
$$= y_{n,\mu\sigma} y^n_{,\nu} + y_{n,\nu\sigma} y^n_{,\mu}, \tag{7.1}$$

since we can move the suffix n freely up and down, on account of the constancy of the h_{mn}. Interchanging μ and σ in (7.1) we get

$$g_{\sigma\nu,\mu} = y_{n,\sigma\mu} y^n_{,\nu} + y_{n,\nu\mu} y^n_{,\sigma}. \tag{7.2}$$

Interchanging ν and σ in (7.1)

$$g_{\mu\sigma,\nu} = y_{n,\mu\nu} y^n_{,\sigma} + y_{n,\sigma\nu} y^n_{,\mu}. \tag{7.3}$$