General Theory of Relativity

P.A.M. Dirac
Florida State University

1975

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

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

so that the four coordinates may be written x^μ, where the suffix μ takes on the values 0, 1, 2, 3. The suffix is written in the upper position so 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,

$$\left(dx^0\right)^2 - \left(dx^1\right)^2 - \left(dx^2\right)^2 - \left(dx^3\right)^2 \qquad (1.1)$$

is invariant.

Any set of 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

$$\left(A^0\right)^2 - \left(A^1\right)^2 - \left(A^2\right)^2 - \left(A^3\right)^2 = (A, A) \qquad (1.2)$$

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

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

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

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

Then the expression on the left 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 equation (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 A^μ, because of the differences in sign, and the vector is called a *covariant vector.*([1])

From 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. Those 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^ν, as distinct of the scalar product (1.3) which is called the inner product.

The tensor $A^\mu B^\nu$ is rather a 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 components of $A^\mu B^\nu$.

[1] In fact, if we set $A_\mu = g_{\mu\nu} A^\nu$ where

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

we can see that the coefficients of the transformation $A'_\mu = \Lambda_\mu{}^\nu A_\nu$ are linearly related to the coefficients of the transformation $A'^\mu = \Lambda^\mu{}_\nu A^\nu$.

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 $T^\mu{}_\nu$.

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 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 vector 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 only downstairs. A suffix appearing 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 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 summation 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 $g_{\mu\nu}$.[2]

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. It's squared length is

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

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

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

since in the second term in the left we can interchange μ and ν and then set $g_{\mu\nu} = g_{\nu\mu}$. 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 the determinant of $g_{\mu\nu}$. It must not vanish; otherwise the four axes would not provide independent directions in space-time and would no 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

[2] In general the number of independent components in a symmetric matrix of $n \times n$ dimensions is equal to the number of unordered pairs of n elements. From the total number of ordered pairs we subtract n pairs with identical elements, and we are left with $n^2 - n$. We divide the last number by two to obtain the number of unordered pairs with distinct elements. The total number of uordered pairs is then

$$T_n = \frac{1}{2} n \cdot (n+1)$$

which is the nth triangular number. Then we notice that T_{n-1} is the number of independent components of an antisymmetric matrix of $n \times n$ dimensions. In the case of four dimensions this number is $T_3 = 6$.

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 $g_{\mu\nu}$, 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 agian, he result is the same as the original tensor, on account of (2.4) and 2.5. Note that g_μ^ν just produces a substitution of ρ for μ or of μ for ρ,

$$g_\mu^\rho A^\mu = A^\rho,$$

and of μ for ρ,

$$g_\mu^\rho A_\rho = A_\mu.$$

if we apply the rule to raising a suffix to the μ in $g_{\mu\rho}$ we get

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

This agrees with (2.4) if we take into accoun 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 form (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 in space. It then becomes a field quantity.

If we take such 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 a 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^{μ}'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 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 he 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

$$\frac{\partial x^\lambda}{\partial x^{\mu'}}\frac{\partial x^{\mu'}}{\partial x^\nu} = g^\lambda_\nu. \tag{3.4}$$

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 $x^\lambda_{,\nu'}$ 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.([3])

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},

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

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

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

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 as a function of the four x^{μ} or 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 equations (3.5) and thus *the derivative of a scalar field is a covariant vector field.*([4])

[3]In fact if $T^{\mu\nu}$ is tensor and ($T^{\mu\nu} = \pm T^{\nu\mu}$). By the transformation law

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

Using the property of symmetry

$$T^{\mu'\nu'} = \pm x^{\mu'}_{,\alpha} x^{\nu'}_{,\beta} T^{\beta\alpha} = \pm x^{\nu'}_{,\beta} x^{\mu'}_{,\alpha} T^{\beta\alpha}$$

and reordering the dummy suffixes

$$T^{\mu'\nu'} = \pm x^{\nu'}_{,\alpha} x^{\mu'}_{,\beta} T^{\alpha\beta} = \pm T^{\nu'\mu'}.$$

[4]As a matter of fact, a covariant (cogradient) vector can be defined as an array of four quantities A_{μ} that transform like the components of the gradient of a scalar. A contravariant (contragradient) vector can be defined as an array of four quantities that transform like the differentials dx^{μ}.

4 Nontensors

We can have a quantity $N^{\mu}_{\nu\rho\cdots}$ 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 a 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 apparing together in the same equation. The rules for balancing suffixes apply equally to tensors and nontensors.

4.1 The Quotient Theorem

Suppose that $P_{\lambda\mu\nu}$ 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 this is a tensor; therefore

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

Thus

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

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 is $P_{\alpha\beta\gamma}$ 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 will 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 immersed in a flat space of a higher number of dimensions; say N. In this N-dimensional space we introduce rectilinear coordinates z^n $(n = 1, \cdots, N)$. Those coordinates do not need to be orthogonal, only rectilinear. Between two neighboring points there is an invariant distance ds given by

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

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

$$dz_n = h_{mn} dz^m.$$

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

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

The squared distance between them is, from (6.1)

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

We may write it

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

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). Thus

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

Now, shift the vector A^n, keeping it parallel to itself (which means, of course, keeping its 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 on 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([5])

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

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

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

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

$$\begin{aligned} K_\nu(x + dx) &= A^n \left[y_{n,\nu}(x) + y_{n,\nu,\sigma} dx^\sigma \right] \\ &= A^\mu y^n_{,\mu} \left[y_{n,\nu} + y_{n,\nu,\sigma} dx^\sigma \right] \\ &= A_\nu + A^\mu y^n_{,\mu} y_{n,\nu,\sigma} dx^\sigma \end{aligned}$$

[5]Here, Dirac assumes that the metric is positive (or negative) which is not the case for the space-time of relativity. Therefore, this mathematical argument supports, at most, an analogy between Einstein's space-time and a Riemann space.

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**)

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

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

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

Interchanging ν and σ in (7.1)

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

Now take (7.1)+(7.3)−(7.2) and divide by 2. The result is

$$\frac{1}{2}\left(g_{\mu\nu,\sigma} + g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu}\right) = y_{n,\nu\sigma} y^n_{,\mu}. \tag{7.4}$$

Put

$$\Gamma_{\mu\nu\sigma} = \frac{1}{2}\left(g_{\mu\nu,\sigma} + g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu}\right) \tag{7.5}$$

It is called a Christoffel symbol of the first kind. It is symmetrical with respect to the last two suffixes. It is a nontensor. A simple consequence of (7.5) is

$$\Gamma_{\mu\nu\sigma} + \Gamma_{\nu\mu\sigma} = g_{\mu\nu,\sigma}. \tag{7.6}$$

We see now that (6.7) can be written as

$$dA_\nu = A^\mu \Gamma_{\mu\nu\sigma} dx^\sigma. \tag{7.7}$$

Al reference to the N-dimensional space has now disappeared, as the Christoffel symbol involves only the metric $g_{\mu\nu}$ of physical space.

We can infer that the length of a vector is unchanged by parallel displacement. We have

$$\begin{aligned}
d\left(g^{\mu\nu}A_\mu A_\nu\right) &= g^{\mu\nu}A_\nu dA_\mu + g^{\mu\nu}A_\mu dA_\nu + A_\mu A_\nu g^{\mu\nu}{}_{,\sigma}dx^\sigma \\
&= A^\mu dA_\mu + A^\nu dA_\nu + A_\mu A_\nu g^{\mu\nu}{}_{,\sigma}dx^\sigma \\
&= A^\mu A^\nu \Gamma_{\nu\mu\sigma}dx^\sigma + A^\nu A^\mu \Gamma_{\mu\nu\sigma}dx^\sigma + A_\mu A_\nu g^{\mu\nu}{}_{,\sigma}dx^\sigma \\
&= A^\mu A^\nu g_{\mu\nu,\sigma}dx^\sigma + A_\mu A_\nu g^{\mu\nu}{}_{,\sigma}dx^\sigma
\end{aligned} \tag{7.8}$$

Now $g^{\alpha\mu}{}_{,\sigma}g_{\mu\nu} + g^{\alpha\mu}g_{\mu\nu,\sigma} = \left(g^{\alpha\mu}g_{\mu\nu}\right)_{,\sigma} = g^\alpha_{\nu,\sigma} = 0$. Multiplying by $g^{\beta\nu}$, we get

$$g^{\alpha\beta}{}_{,\sigma} = -g^{\alpha\mu}g^{\beta\nu}g_{\mu\nu,\sigma}. \tag{7.9}$$

This is a useful formula giving the derivative of $g^{\alpha\beta}$ in terms of the derivative of $g_{\mu\nu}$. It allows us to infer

$$A_\alpha A_\beta g^{\alpha\beta}{}_{,\sigma} = -A^\mu A^\nu g_{\mu\nu,\sigma}$$

and so the expression (7.8) vanishes. Thus the length of a vector is constant. In particular, a null vector (i.e. a vector of zero length) remains a null vector under parallel displacement.

The constancy of the length of the vector follows also from geometrical arguments. When we split up the vector A^μ into tangential and normal parts according to (6.5), the normal part is infinitesimal and is orthogonal to the tangential part. It follows that, to the first order, the length of the whole vector equals that of its tangential part.

The constancy of the length of any vector requires the constancy of the scalar product $g^{\mu\nu}A_\mu B_\nu$ of any two vectors A and B. This can be inferred from the constancy of the length of $A + \lambda B$ for any value of the parameter λ.

It is frequently useful to raise the first suffix of the Christoffel symbol so as to form

$$\Gamma^\mu_{\nu\sigma} = g^{\mu\lambda}\Gamma_{\lambda\mu\sigma}.$$

It is then called a Christoffel symbol of the second kind. It is symmetrical between its two lower suffixes. As explained in Section 4, this raising is quite permissible, even for a nontensor.

The formula (7.7) may be written

$$dA_\nu = \Gamma^\mu_{\nu\sigma}A_\mu dx^\sigma \tag{7.10}$$

It is the standard formula referring to covariant components. For a second vector B^ν we have

$$\begin{aligned} d(A_\nu B^\nu) &= 0 \\ A_\nu dB^\nu &= -B^\nu dA_\nu = -B^\nu \Gamma^\mu_{\nu\sigma} A_\mu dx^\sigma \\ &= -B^\mu \Gamma^\nu_{\mu\sigma} A_\nu dx^\sigma \end{aligned}$$

This must hold for any A_ν, so we get

$$dB^\nu = -\Gamma^\nu_{\mu\sigma} B^\mu dx^\sigma. \tag{7.11}$$

This is the standard formula for parallel displacement referring to contravariant components.

8 Geodesics

Take a point with coordinates z^μ and suppose it moves along a track; we then have a function of some parameter τ. Put $dz^\mu/d\tau = u^\mu$.

There is a vector u^μ at each point of the track. Suppose that as we go along the track the vector u^μ gets shifted by parallel displacement. Then the whole track is determined if we are given the initial point and the initial value of the vector u^μ. We just have to shift the initial point from z^μ to $z^\mu + u^\mu d\tau$, then shift the vector u^μ to this new point by parallel displacement, then shift the point again in the direction fixed by the new u^μ, and so on. Not only is the track determined, but also the parameter τ along it. A track produced this way is called a geodesic.

If the vector u^μ is a null vector, it always remains a null vector and the track is called a null geodesic. If the vector u^μ is initially timelike (i.e., $u^\mu u_\mu > 0$), it is always timelike and we have a timelike geodesic. Similarly if u^μ is initially spacelike ($u^\mu u_\mu < 0$), it is always spacelike and we have a spacelike geodesic.

We get the equations of a geodesic by applying (7.11) with $B^\nu = u^\nu$ and $dx^\sigma = dz^\sigma$. Thus

$$\frac{du^\nu}{d\tau} + \Gamma^\nu_{\mu\sigma} u^\mu \frac{dz^\sigma}{d\tau} = 0 \tag{8.1}$$

or

$$\frac{d^2 z}{d\tau^2} = +\Gamma^\nu_{\mu\sigma} \frac{dz^\mu}{d\tau} \frac{dz^\sigma}{d\tau} = 0 \tag{8.2}$$

For a timelike geodesic we may multiply the initial u^ν by a factor so as to make its length unity. This merely requires a change in the scale of τ.

The vector u^μ now always has the length unity. It is just the velocity vector $v^\mu = dz^\mu/ds$, and the parameter τ has become the proper time s.

Equation (8.1) then becomes

$$\frac{dv^\mu}{ds} + \Gamma^\mu_{\nu\sigma} u^\nu u^\sigma = 0. \tag{8.3}$$

Equation (8.2) becomes

$$\frac{d^2 z^\mu}{ds^2} + \Gamma^\mu_{\nu\sigma} \frac{dz^\nu}{ds}\frac{dz^\sigma}{ds} = 0. \tag{8.4}$$

We make the physical assumption that the world line of a particle not acted on by any forces, except gravitational, is a timelike geodesic. This replaces Newton's first law of motion. Equation (8.4) fixes the acceleration and provides the equations of motion.

We also make the assumption that the path of a ray of light is a null geodesic. It is fixed by equation (8.2) referring to some parameter along the path. The proper time s cannot now be used because ds vanishes.

9 The stationary property of geodesics

A gepdesic that is not a null geodesic has the property that $\int ds$, taken along a sectio of the track with the end points P and Q, is stationary if one makes a small variation of the tack keeping the end points fixed.

Let us suppose that each point of the track, with the coordinates z^μ is shifted so that its coordinates become $z^\mu + \delta z^\mu$. If dx^μ denotes an element along the track,

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

Thus

$$\begin{aligned} 2ds\delta ds &= dx^\mu dx^\nu \delta g_{\mu\nu} + g_{\mu\nu} dx^\mu \delta dx^\nu + g_{\mu\nu} dx^\nu \delta dx^\mu \\ &= dx^\mu dx^\nu \delta g_{\mu\nu} + 2 g_{\mu\lambda} dx^\mu \delta dx^\lambda. \end{aligned}$$

Now

$$\delta dx^\lambda = d(\delta x^\lambda).$$

Thus, with the help of $dx^\mu = v^\mu ds$,

$$\delta(ds) = \left(\frac{1}{2} g_{\mu\nu,\lambda} v^\mu v^\nu \delta x^\lambda + g_{\mu\lambda} v^\mu \frac{d\delta x^\lambda}{ds} \right) ds$$

Hence

$$\delta \int ds = \int \delta(ds) = \int \left[\frac{1}{2} g_{\mu\nu,\lambda} v^\mu v^\nu \delta x^\lambda + g_{\mu\lambda} v^\mu \frac{d\delta x^\lambda}{ds} \right] ds$$

By partial integration (of the second term), using the condition that $\delta x^\lambda = 0$ at the end points P and Q, we get

$$\delta \int ds = \int \left[\frac{1}{2} g_{\mu\nu,\lambda} v^\mu v^\nu - \frac{d}{ds} \left(g_{\mu\lambda} v^\mu \right) \right] \delta x^\lambda ds \tag{9.1}$$

The condition for this to vanish with arbitrary δx^λ is

$$\frac{d}{ds} \left(g_{\mu\lambda} v^\mu \right) - \frac{1}{2} g_{\mu\nu,\lambda} = 0 \tag{9.2}$$

Now

$$\begin{aligned} \tfrac{d}{ds} \left(g_{\mu\lambda} v^\mu \right) &= g_{\mu\lambda} \tfrac{dv^\mu}{ds} + g_{\mu\lambda,\nu} v^\mu v^\nu \\ &= g_{\mu\lambda} \tfrac{dv^\mu}{ds} + \tfrac{1}{2} \left(g_{\lambda\mu,\nu} + g_{\lambda\nu,\mu} \right) v^\mu v^\nu. \end{aligned}$$

Thus the condition (9.2) becomes

$$g_{\mu\lambda} \frac{dv^\mu}{ds} + \Gamma_{\lambda\mu\nu} v^\mu v^\nu = 0.$$

Multiplying this by $g^{\lambda\sigma}$, it becomes

$$\frac{dv^\sigma}{ds} + \Gamma^\sigma_{\mu\nu} v^\mu v^\nu = 0,$$

which is just the condition (8.3) for the geodesic.

This work shows that for a geodesic, (9.1) vanishes and $\int ds$ is stationary. Conversely, if we assume that $\int ds$ is stationary, we can infer that the track is a geodesic. Thus we may use the stationary condition as the definition of a geodesic, except in the case of a null geodesic.

10 Covariant differentiation

Let S be a scalar field. Its derivative S, y is a covariant vector, as we saw in Section 3. Now let A_μ be a vector field. Is its derivative $A_{\mu,\nu}$ a tensor?

We must examine how $A_{\nu,\nu}$ transforms under a change of coordinate system. With the notation in Section 3, A_μ transforms to

$$A_{\mu'} = A_\rho x^\rho_{,\mu'}$$

like equation (3.5), and hence

$$\begin{aligned} A_{\mu',\nu'} &= \left(A_\rho x^\rho_{,\mu'}\right)_{,\nu'} \\ &= A_{\rho,\sigma} x^\sigma_{,\nu'} x^\rho_{,\mu'} + A_\rho x^\rho_{,\mu'\nu'}. \end{aligned}$$

The last term should not be here if we were to have the correct transformation law for a tensor. Thus $A_{\mu,\nu}$ is not a tensor.

We can, however, modify the process of differentiation so as to get a tensor. Let us take the vector A_μ at the point x and shift it to $x + dx$ by parallel displacement. It is still a vector. We may subtract it from the vector A_μ at $x + dx$ and the difference will be a vector. It is, to the first order

$$A_\mu(x + dx) - \left[A_\mu(x) + \Gamma^\alpha_{\mu\nu} A_\alpha dx^\nu\right] = \left(A_{\mu,\nu} - \Gamma^\alpha_{\mu\nu} A_\alpha\right) dx^\nu.$$

This quantity is a vector, for any vector dx^ν; hence, by the quotient theorem of Section 4, the coefficient

$$A_{\mu\nu} - \Gamma^\alpha_{\mu\nu} A_\alpha$$

is a tensor. One can easily verify directly that it transforms correctly under a change of coordinate system.

It is called the covariant derivative of A_μ and written

$$A_{\mu:\nu} = A_{\mu,\nu} - \Gamma^\alpha_{\mu\nu} A_\alpha. \tag{10.1}$$

The sign : before a lower suffix will always denote a covariant derivative, just as the comma denotes an ordinary derivative.

Let B_ν be a second vector. We define the outer product $A_\mu B_\nu$ to have the covariant derivative([6])

$$\left(A_\mu B_\nu\right)_\sigma = A_{\mu:\sigma} B_\nu + A_\mu B_{\nu:\sigma} \tag{10.2}$$

Evidently it is a tensor with three suffixes. It has the value

$$\left(A_\mu B_\nu\right)_\sigma = \left(A_{\mu,\sigma} - \Gamma^\alpha_{\mu\sigma} A_\alpha\right) B_\nu + A_\mu \left(B_{\nu,\sigma} - \Gamma^\alpha_{\nu\sigma} B_\alpha\right)$$

[6]Notice that the definition of the covariant derivative of a second rank tensor is not motivated because the notion of *parallel displacement of tensor* doesn't make sense.

Let $T_{\mu\nu}$ be a tensor with two suffixes. It is expressible as a sum of terms like $A_\mu B_\nu$, so, its covariant derivative is

$$T_{\mu\nu:\sigma} = T_{\mu\nu,\sigma} - \Gamma^\alpha_{\mu\sigma} T_{\alpha\nu} - \Gamma^\alpha_{\nu\sigma} T_{\mu\alpha}. \tag{10.3}$$

The rule can be extended to the covariant derivative of a tensor $Y_{\mu\nu\cdots}$ with any number of suffixes downstairs

$$Y_{\mu\nu\cdots:\sigma} = Y_{\mu\nu\cdots,\sigma} - \text{ a } \Gamma \text{ term for each suffix.} \tag{10.4}$$

In each of these Γ terms we must make the suffixes balance, which is sufficient to fix how the suffixes go.

The case of a scalar is included in the general formula (10.4) with the number of suffixes in Y zero.

$$Y_{:\sigma} = Y_{,\sigma}. \tag{10.5}$$

Let's apply (10.3) to the fundamental tensor $g_{\mu\nu}$. It gives

$$\begin{aligned} g_{\mu\nu:\sigma} &= g_{\mu\nu,\sigma} - \Gamma^\alpha_{\mu\sigma} g_{\alpha,\nu} - \Gamma^\alpha_{\nu\sigma} g_{\mu\alpha} \\ &= g_{\mu\nu,\sigma} - \Gamma_{\nu\mu\sigma} - \Gamma_{\nu\mu\sigma} = 0 \end{aligned}$$

from (7.6). Thus the $g_{\mu\nu}$, count as constants under covariant differentiation.

Formula (10.2) is the usual rule that one uses for differentiating a product. We assume this usual rule holds also for the covariant derivative of the scalar product of two vectors. Thus

$$(A^\mu B_\mu) : \sigma = A^\mu_{:\sigma} B_\mu + A^\mu B_{\mu:\sigma}.$$

We get, according to (10.5) and (10.1),

$$(A^\mu B_\mu) : \sigma = A^\mu_{:\sigma} B_\mu + A^\mu \left(B_{\mu,\sigma} - \Gamma^\alpha_{\mu\sigma} B_\alpha \right).$$

and hence

$$A^\mu_{,\sigma} B_\mu = A^\mu_{:\sigma} B_\mu - A^\alpha \Gamma^\mu_{\alpha\sigma} B_\mu.$$

Since this holds for any B_μ, we get

$$A^\mu_{:\sigma} B_\mu = A^\mu_{,\sigma} B_\mu + A^\alpha \Gamma^\mu_{\alpha\sigma} B_\mu, \tag{10.6}$$

which is the basic formula for the covariant derivative of a contravariant vector. The same Christoffel symbol occurs as in the basic formula (10.1) for a covariant vector, but now there is a + sign. The arrangement of suffixes is completely determined by the balancing requirement.

We can extend the formalism so as to include the covariant derivative of any tensor with any number of upstairs or downstairs suffixes. A Γ terms appears for each suffix, with a + sign if the suffix is upstairs an a $-$ sign if it is downstairs. If we contract two suffixes in the tensor, the corresponding Γ terms cancel.

The formula for the covariant derivative of a product,

$$(XY)_{:\sigma} = X_{:\sigma} Y + X Y_{:\sigma}, \tag{10.7}$$

holds quite generally, with X and Y any kind of tensor quantities. On account of the $g_{\mu\nu}$ counting as constants, we can shift suffixes up or down before covariant differentiation and the result is the same as if we shifted them afterwards.

The covariant derivative of a nontensor has no meaning.

The laws of physics must be valid in all systems of coordinates. Thus must be expressible as tensor equations. Whenever they involve the derivative of a field quantity, it must be a covariant derivative. The field equations of physics must all be rewritten with the ordinary derivatives replaced by covariant derivatives. For example, the d'Alembert equation $\Box V = 0$ for a scalar V becomes, in covariant form

$$g^{\mu\nu} V_{:\mu:\nu} = 0.$$

This gives, from (10.1) and (10.5),

$$g^{\mu\nu} \left(V_{,\mu\nu} - \Gamma^{\alpha}_{\mu\nu} V_{\alpha} = 0. \right) \tag{10.8}$$

Even if one is working with flat space (which means neglecting the gravitational field) and one is using curvilinear coordinates, one must write one's equations in terms of covariant derivatives if one wants them to hold in all systems of coordinates.

11 The curvature tensor

With the product law (10.7) we see that covariant differentiation is very similar to ordinary differentiation. But there is an important property of ordinary differentiation, that if we perform two differentiations in succession their order does not matter, which does not, in general, hold for covariant differentiation.

Let us first consider a scalar field S. We have from the formula (10.1),

$$\begin{aligned} S_{:\mu:\nu} &= S_{:\mu,\nu} - \Gamma^{\alpha}_{\mu\nu} S_{:\alpha} \\ &= S_{,\mu\nu} - \Gamma^{\alpha}_{\mu\nu} S_{,\alpha}. \end{aligned} \tag{11.1}$$

This is symmetrical between μ and ν, so in this case the order of the covariant differentiation does not matter.

Now let us take a vector A_ν and apply two covariant differentiations to it. From the formula (10.3) with $A_{\nu:\rho}$ for $T_{\nu\rho}$ we get

$$\begin{aligned} A_{\nu:\rho:\sigma} &= A_{\nu:\rho,\sigma} - \Gamma^{\alpha}_{\nu\sigma} A_{\alpha:\rho} - \Gamma^{\alpha}_{\rho\sigma} A_{\nu:\alpha} \\ &= \left(A_{\nu,\rho} - \Gamma^{\alpha}_{\nu\rho} A_{\alpha}\right)_{,\sigma} - \Gamma^{\alpha}_{\nu\sigma}\left(A_{\alpha,\rho} - \Gamma^{\beta}_{\alpha\rho} A_{\beta}\right) - \Gamma^{\alpha}_{\rho\sigma}\left(A_{\nu,\alpha} - \Gamma^{\beta}_{\nu\alpha} A_{\beta}\right) \\ &= A_{\nu,\rho\sigma} - \Gamma^{\alpha}_{\nu\rho} A_{\alpha,\sigma} - \Gamma^{\alpha}_{\nu\sigma} A_{\alpha,\rho} - \Gamma^{\alpha}_{\rho\sigma} A_{\nu,\alpha} \\ &\quad - A_{\beta}\left(\Gamma^{\beta}_{\nu\rho,\sigma} - \Gamma^{\alpha}_{\nu\sigma}\Gamma^{\beta}_{\alpha\rho} - \Gamma^{\alpha}_{\rho\sigma}\Gamma^{\beta}_{\nu\alpha}\right) \end{aligned}$$

Interchange ρ and σ here and subtract from the previous expression. The result is

$$A_{\nu:\rho:\sigma} - A_{\nu:\sigma:\rho} = A_{\beta} R^{\beta}_{\nu\rho\sigma} \tag{11.2}$$

where

$$R^{\beta}_{\nu\rho\sigma} = \Gamma^{\beta}_{\nu\sigma,\rho} - \Gamma^{\beta}_{\nu\rho,\sigma} + \Gamma^{\alpha}_{\nu\sigma}\Gamma^{\beta}_{\alpha\rho} - \Gamma^{\alpha}_{\nu\rho}\Gamma^{\beta}_{\alpha\sigma} \tag{11.3}$$

The left hand of (11.2) is a tensor. It follows that the right-hand side of (11.2) is a tensor. This holds for any vector A_β; therefore, by the quotient theorem in Section 4, $R^{\beta}_{\nu\rho\sigma}$ is a tensor. It is called the Riemann-Christoffel tensor or the curvature tensor.

It has the obvious property

$$R^{\beta}_{\nu\rho\sigma} = -R^{\beta}_{\nu\sigma\rho} \tag{11.4}$$

Also we easily see from (11.3) that

$$R^{\beta}_{\nu\rho\sigma} + R^{\beta}_{\rho\sigma\nu} + R^{\beta}_{\sigma\nu\rho} = 0 \tag{11.5}$$

Let us lower the suffix β and put it as the first suffix. We get

$$R_{\mu\nu\sigma\rho} = g_{\mu\beta} R^{\beta}_{\nu\sigma\rho} = g_{\mu\beta}\Gamma^{\beta}_{\nu\sigma,\rho} + \Gamma^{\alpha}_{\nu\alpha}\Gamma_{\mu\alpha\rho} - \langle\rho\sigma\rangle,$$

where the symbol $\langle\rho\sigma\rangle$ is used to denote the preceding terms with ρ and σ interchanged. Thus

$$\begin{aligned} R_{\mu\nu\sigma\rho} &= R_{\mu\nu\sigma,\rho} - g_{\mu\beta,\rho}\Gamma^{\beta}_{\nu\sigma} + \Gamma_{\mu\beta\rho}\Gamma^{\beta}_{\nu\sigma} - \langle\rho\sigma\rangle \\ &= \Gamma_{\mu\nu\sigma,\rho} - \Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma} - \langle\rho\sigma\rangle \end{aligned}$$

from (7.6). So from (7.5)

$$R_{\mu\nu\sigma\rho} = \frac{1}{2}\left(g_{\mu\sigma,\nu\rho} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma} - g_{\nu\rho,\nu\sigma}\right) + \Gamma_{\beta\mu\sigma}\Gamma^{\beta}_{\nu\rho} - \Gamma_{\beta\mu\rho}\Gamma^{\beta}_{\nu\sigma} \tag{11.6}$$

Some further symmetries now show up; namely,

$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma} \tag{11.7}$$

and

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu} = R_{\sigma\rho\nu\mu} \tag{11.8}$$

The result of all these symmetries is that, of the 256 components of $R_{\mu\nu\rho\sigma}$, only 20 are independent.

12 The condition for flat Space

If space is flat, we may choose a system of coordinates that is rectilinear and then the $g_{\mu\nu}$ are constant. The tensor $R_{\mu\nu\rho\sigma}$ then vanishes.

Conversely, if $R_{\mu\nu\rho\sigma}$ vanishes, one can prove that the space is flat. Let us take a vector A_μ situated at the point x and shift it by parallel displacement to the point $x + dx$. Then shift it by parallel displacement to the point $x + dx + \delta x$. If $R_{\mu\nu\rho\sigma}$ vanishes, he result must be the same as if we had shifted it first from x to $x + \delta x$, then to $x + \delta x + dx$. Thus we can shift the vector to a distant point and the result we get is independent of the path to the distant point. Therefore, if we shift the original vector A_μ at x to all points by parallel displacement, we get a vector field that satisfies $A_{\mu:\nu} = 0$, or

$$A_{\mu,\nu} = \Gamma^{\alpha}_{\mu\nu}. \tag{12.1}$$

Can such a vector field be the gradient of a scalar? Let us put $A_\mu = S_\mu$ in (12.1). We get

$$S_{,\mu\nu} = \Gamma^{\alpha}_{\mu\nu}S_\alpha \tag{12.2}$$

By virtue of the symmetry of $\Gamma^\alpha_{\mu\nu}$ in the lower suffixes, we have the same value for $S_{,\mu\nu}$ as $S_{,\nu\mu}$ and the equations (12.2) are integrable.

Let us take four independent scalars satisfying (12.2) and let them to be the coordinates $x^{\alpha'}$ of a new system of coordinates. Then

$$x^{\alpha'}_{,\mu\nu} = \Gamma^\sigma_{\mu\nu} x^{\alpha'}_{,\sigma}.$$

According the transformation law (3.7),

$$g_{\mu\lambda} = g_{\alpha'\beta'} x^{\alpha'}_{,\mu} x^{\beta'}_{,\lambda}.$$

Differentiating this equation with respect to x^ν, we get

$$\begin{aligned} g_{\mu\lambda,\nu} - g_{\alpha'\beta',\nu} x^{\alpha'}_{,\mu} x^{\beta'}_{,\lambda} &= g_{\alpha'\beta'} \left(x^{\alpha'}_{,\mu\nu} x^{\beta'}_{,\lambda} + x^{\alpha'}_{,\mu} x^{\beta'}_{,\lambda\nu} \right) \\ &= g_{\alpha'\beta'} \left(\Gamma^\sigma_{\mu\nu} x^{\alpha'}_{,\sigma} x^{\beta'}_{,\lambda} + x^{\alpha'}_{,\mu} \Gamma^\sigma_{\lambda\nu} x^{\beta'}_{,\sigma} \right) \\ &= g_{\sigma\lambda} \Gamma^\sigma_{\mu\nu} + g_{\mu\sigma} \Gamma^\sigma_{\lambda\nu} \\ &= \Gamma_{\lambda\mu\nu} + \Gamma_{\mu\lambda\nu} = g_{\mu\lambda,\nu} \end{aligned}$$

from (7.6). Thus

$$g_{\alpha'\beta',\nu} x^{\alpha'}_{,\mu} x^{\beta'}_{,\lambda} = 0.$$

It follows that $g_{\alpha'\beta',\nu} = 0$. Referred to the new system of coordinates, the fundamental tensor is constant. Thus we have flat space referred to rectilinear coordinates.

13 The Bianci relations

To deal with the second covariant derivative of a tensor, take first the case in which the tensor is the outer product of two vectors $A_\mu B_\tau$. We have

$$\begin{aligned} (A_\mu B_\tau)_{:\rho:\sigma} &= (A_{\mu:\rho} B_\tau + A_\mu B_{\tau:\rho}) \\ &= A_{\mu:\rho:\sigma} B_\tau + A_{\mu:\rho} B_{\tau:\sigma} + A_{\mu:\sigma} B_{\tau:\rho} + A_\mu B_{\tau:\rho:\sigma}. \end{aligned}$$

Now interchange ρ and σ and subtract. We get from (11.2)

$$(A_\mu B_\tau)_{:\rho:\sigma} - (A_\mu B_\tau)_{:\sigma:\rho} = A_\alpha R^\alpha_{\mu\rho\alpha} B_\tau + A_\mu R^\alpha_{\tau\rho\sigma} B_\alpha.$$

A general tensor $T_{\mu\tau}$ is expressible as a sum of terms like $A_\mu B_\tau$, so it must satisfy

$$T_{\mu\tau:\rho:\sigma} - T_{\mu\tau:\sigma:\rho} = T_{\alpha\tau} R^\alpha_{\mu\rho\sigma} - T_{\mu\alpha} R^\alpha_{\tau\rho\sigma}. \tag{13.1}$$

Now take $T_{\mu\tau}$ to be the covariant derivative of a vector $A_{\mu:\tau}$. We get

$$A_{\mu:\tau:\rho:\sigma} - A_{\mu:\tau:\sigma:\rho} = A_{\alpha:\tau} R^{\alpha}_{\mu\rho\sigma} + A_{\mu:\alpha} R^{\alpha}_{\tau\rho\sigma}.$$

In this formula make cyclic permutations of τ, ρ, σ and add the three equations so obtained. The left hand side gives

$$\begin{aligned} & A_{\mu:\tau:\rho:\sigma} - A_{\mu:\tau:\sigma:\rho} + \text{ cyc perm} \\ = \; & \left(A_{\alpha} R^{\alpha}_{\mu\rho\sigma}\right)_{:\tau} + \text{ cyc perm} \\ = \; & A_{\alpha:\tau} R^{\alpha}_{\mu\rho\sigma} + A_{\mu:\alpha} R^{\alpha}_{\tau\rho\sigma} + \text{ cyc perm.} \end{aligned} \tag{13.2}$$

The right-hand side gives

$$A_{\alpha:\tau} R^{\alpha}_{\mu\rho\sigma} + \text{ cyc perm,} \tag{13.3}$$

as the remaining terms cancel from (11.5). The first term of (13.2) cancels with (13.3) and we are left with

$$A_{\alpha} R^{\alpha}_{\mu\rho\sigma:\tau} + \text{ cyc perm} = 0.$$

The factor A_{α} occurs throughout this equation and may be canceled out. We are left with

$$R^{\alpha}_{\mu\rho\sigma:\tau} + R^{\alpha}_{\mu\sigma\tau:\rho} + R^{\alpha}_{\mu\tau\rho:\sigma} = 0 \tag{13.4}$$

The curvature tensor satisfies these differential equations as well as other symmetry relations in Section 11. They are known as the Bianci relations.

14 The Ricci tensor

Let us contract two of the suffixes in $R_{\mu\nu\rho\sigma}$. If we take two with respect to which it is antisymmetrical, we get zero, of course. If we take any other two we get the same result, apart from the sign because of the symmetries (11.4), (11.7), and (11.8). Let us take the first and last and put

$$R^{\alpha}_{\nu\rho\alpha} = R_{\nu\rho}.$$

It is called the Ricci tensor.

By multiplying (11.8) by $g^{\mu\nu}$ we get

$$R_{\nu\rho} = R_{\rho\nu}. \tag{14.1}$$

The Ricci tensor is symmetrical.

We may contract again and form

$$g^{\nu\rho} R_{\nu\rho} = R^{\nu}_{\nu} = R,$$

say. This R is a scalar and is called the scalar curvature or total curvature. It is defined in such way that it is positive for the surface of a sphere in three dimensions, as one can check by a straightforward calculation.

The Bianci relation (13.4) involves five suffixes. Let us contract it twice and get a relation with one nondummy suffix. Put $\tau = \alpha$ and multiply by $g^{\mu\rho}$. The result is

$$g^{\mu\rho}\left(R^{\alpha}_{\mu\rho\sigma:\alpha} + R^{\alpha}_{\mu\sigma\alpha:\rho} + R^{\alpha}_{\mu\alpha\rho:\sigma}\right) = 0$$

or

$$\left(g^{\mu\rho} R^{\alpha}_{\mu\rho\sigma}\right)_{:\alpha} + \left(g^{\mu\rho} R^{\alpha}_{\mu\sigma\alpha}\right)_{:\rho} + \left(g^{\mu\rho} R^{\alpha}_{\mu\alpha\rho}\right)_{:\sigma} = 0. \tag{14.2}$$

Now

$$g^{\mu\rho} R^{\alpha}_{\mu\rho\sigma} = g^{\mu\rho} g^{\alpha\beta} R_{\beta\mu\rho\sigma} = g^{\mu\rho} g^{\alpha\beta} R_{\mu\beta\sigma\rho} = g^{\alpha\beta} R_{\beta\sigma} = R^{\alpha}_{\sigma}.$$

One can write R^{α}_{σ} with the suffixes one over the other on account of $R_{\alpha\sigma}$ being symmetrical. Equation (14.2) now becomes

$$R^{\alpha}_{\sigma:\alpha} + \left(g^{\mu\rho} R_{\mu\sigma}\right)_{:\rho} - R_{:\sigma} = 0$$

or

$$2R^{\alpha}_{\sigma:\alpha} - R_{:\sigma} = 0$$

which is the Bianci relation for the Ricci tensor. If we raise the suffix σ, we get

$$\left(R^{\sigma\alpha} - \frac{1}{2} g^{\sigma\alpha} R\right)_{:\alpha} = 0 \tag{14.3}$$

The explicit expression for the Ricci tensor is, from (11.3)

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} \tag{14.4}$$

The first term here does not appear to be symmetrical in μ and ν, although the other three terms evidently are. To establish that the first term really is symmetrical we need a little calculation.

To differentiate the determinant g we must differentiate each element $g_{\lambda\mu}$ and then multiply for the cofactor $gg^{\lambda\mu}$. Thus

$$g_{,\nu} = gg^{\lambda\mu} g_{\lambda\mu,\nu}. \tag{14.5}$$

Hence

$$\begin{aligned}\Gamma^{\mu}_{\lambda\mu} &= g^{\lambda\mu}\Gamma_{\lambda\nu\mu} = \tfrac{1}{2}\left(g_{\lambda\nu,\mu} + g_{\lambda\mu,\nu} - g_{\mu\nu,\lambda}\right) \\ &= \tfrac{1}{2}g^{\lambda\mu}g_{\lambda\mu,\nu} = \tfrac{1}{2}g^{-1}g_{\nu} = \tfrac{1}{2}\left(log(g)\right)_{,\nu}.\end{aligned} \tag{14.6}$$

This makes it evident that the first term of (14.4) is symmetrical.

15 Einstein's Law of gravitation

Up to the present our work has all been pure mathematics (apart from the physical assumption that the track of a particle is a geodesic). It was done mainly in the last century and applies to curved spaces in any number of dimensions. The only place where the number of dimensions would appear in the formalism is in the equation

$$g^{\mu}_{\nu} = \text{ number of dimensions}$$

Einstein made the assumption that in empty spaces

$$R_{\mu\nu} = 0. \tag{15.1}$$

It constitutes his law of gravitation. "Empty" here means that there is no matter present and no physical fields except the gravitational field. The gravitational field does not disturb the emptyness. Other fields do.([7]). The conditions for empty space hold in a good approximation for the space between the planets in the solar system, and equation (15.1) applies there.

Flat space obviously satisfies (15.1). The geodesics are then straight lines and so particles move along straight lines. Where space is not flat, Einstein's law puts restrictions on the curvature. Combined with the assumption that the planets move along geodesics, it gives some information about their motion.

At first sight, Einstein's law of gravitation does not look anything like Newton's. To see a similarity, we must look on the $g_{\mu\nu}$ as *potentials* describing the gravitational field. There are ten of them, instead of just the one potential of the Newtonian theory. They describe not only the gravitational field, but also the system of coordinates. The gravitational field and the system of

[7] If the gravitational field does not disturb the *emptyness* of space, then the gravitational field is nothing, which begs the questions: "What is a quantum of *nothing¿*", "How *nothing* can be the substrate of gravitational waves?", etcetera

coordinates are inextricable mixed up in Einstein's theory, and one cannot describe one without the other.

Looking upon the $g_{\mu\nu}$ as potentials, we find that (15.1) appears as field equations. They are like the usual equations of physics in that they are of the second order, because second order derivatives appear in (14.4), as the Christoffel symbols involve first derivatives. They are unlike the usual field equations in that they are not linear; far from it. The nonlinearity means that the equations are complicated and it is difficult to get accurate solutions.

16 The Newtonian approximation

Let us consider a static gravitational field and refer it to a static coordinate system. The $g_{\mu\nu}$ are then constant in time, $g_{\mu\nu,0} = 0$. Further, we must have

$$g_{m0} = 0,\ (m = 1, \cdots, 3)\,.$$

This leads to

$$g^{m0} = 0,\ g^{00} = (g_{00})^{-1},$$

and g^{mn} is the reciprocal matrix to g_{mn}. Roman suffixes like m and n always take on the values 1, 2, 3. We find that $\Gamma_{m0n} = 0$ and hence also $\Gamma^m_{0n} = 0$.

Let us take a particle that is moving slowly, compared with the velocity of light. Then v^m is a small quantity, of the first order([8]). With neglect of second order quantities,

$$g_{00}{v^0}^2 = 1. \tag{16.1}$$

The particle will move along a geodesic. With neglect of second-order quantities, the equation (8.3) gives

$$\begin{aligned} \tfrac{dv^m}{ds} &= -\Gamma^m_{00}{v^0}^2 = -g^{mn}\Gamma_{n00}{v^0}^2 \\ &= = \tfrac{1}{2}g^{mn}g_{00,n}{v^0}^2. \end{aligned}$$

Now

$$\frac{dv^m}{ds} = \frac{dv^m}{dx^0}\frac{dx^0}{ds} = fracdv^m dx^0 v^0$$

to the first order. Thus

$$\frac{dv^m}{dx^0} = \frac{1}{2}g^{mn}g_{00,n}v^0 = g^{mn}\left({g_{00}}^{1/2}\right)_{,n} \tag{16.2}$$

[8]compared to v^0.

with the help of (16.1). Since the $g_{\mu\nu}$ are independent of x^0, we may lower the suffix m here and get

$$\frac{dv^m}{dx^0} = \left({g_{00}}^{1/2}\right)_{,m}. \tag{16.3}$$

We see that the particle moves as thought it were under the influence of a potential $\left({g_{00}}^{1/2}\right)$. We have not used Einstein's law to obtain this result. We now use Einstein's law to obtain a condition for the potential, so that it can be compared with Newton's.

Let us suppose that the gravitational field is weak, so that the curvature of space is small. Then we may choose our coordinate system so that the curvature of the coordinate lines (each with three x's constant) is small. Under those conditions the $g_{\mu\nu}$ are approximately constant, and $g_{\mu\nu,\sigma}$ and all the Christoffel symbols are small. If we count them of the first order and neglect second-order quantities, Einstein's law (15.1) becomes, from (14.4)

$$\Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} = 0.$$

We can evaluate this most conveniently by contracting (11.6) with ρ and μ interchanged and neglecting second-order terms. The result is

$$g^{\rho\sigma}\left(g_{\rho\sigma,\mu\nu} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma} + g_{\mu\nu,\rho\sigma}\right) = 0. \tag{16.4}$$

Now take $\mu = \nu = 0$ and use the condition that the $g_{\mu\nu}$ are independent of x^0. We get

$$g^{mn} g_{00,mn} = 0. \tag{16.5}$$

The D'Alembert equation (10.9) becomes, in the weak field approximation,

$$g^{\mu\mu} V, \mu\nu = 0.$$

In the static case, this reduces to Laplace's equation

$$g^{mn} V_{,mn} = 0.$$

We may choose our unit of time so that g_{00} is approximately unity. Then we may put

$$g_{00} = 1 + 2V, \tag{16.6}$$

with V small. We get ${g_{00}}^{1/2} = 1 + V$ and V becomes the potential. It satisfies the Laplace equation so that it can be identified with the Newtonian potential, equal to $-m/r$ for a mass m at the origin. To check the sign we see that (16.2) leads to

$$\text{acceleration} = -\text{ grad } V,$$

since g^{mn} has the diagonal elements approximately -1.

We see that Einstein's law of gravitation goes over to Newton's when the field is weak and when it is static.([9]) The successes of the Newtonian theory in explaining the motions of the planets can this be preserved. The static approximation is a good one because the velocities of the planets are all small compared to the with the velocity of light ([10]). The weak field approximation is a good one because the space is very nearly flat. Let us consider some orders of magnitude.

The value of V on the surface of earth turns out to be of the order of 10^{-9}, thus g_{00} given by (16.6) is very close to one. Even so, its difference from one is big enough to produce the important gravitational effects that we see on earth. Taking the radius of earth to be of the order of 10^9 centimeters, we find that $g_{00,m}$ is of the order of 10^{-18}cm^{-1}. The departure from flatness is thus extremely small. However, this has to be multiplied by the square of the speed of light, namely $9 \times 10^{10}\text{cm/sec}^2$, to give the acceleration of gravity at earth's surface. Thus, this acceleration, about 10^3cm/sec^2, is quite appreciable, even thought the department of flatness is far too small to be observed directly.

17 The gravitational red shift

Let us take again a static gravitational field and consider an atom at rest emitting monochromatic radiation. The wavelength of the light will correspond to a definite Δs. Since the atom is at rest we have, for a static system of coordinates such as we used in Section 16,

$$\Delta s^2 = g_{00} \Delta x^{0^2},$$

[9]This is not exactly a proof. From the requirement of covariance and the realization that Einstein's equation are a quasilinear system of second order partial equations, considering that Laplace's operator is the only differential linear operator of the second-degree that satisfies the requirements of homogeneity and isotropy of space, this result is a formal property of the theory. A proof that Einstein's theory reduces to Newton's theory will require to prove that it leads to Poison's equation.

[10]In the local system of reference.

where Δx^0 is the period, that is, the time between successive crests referred to our static coordinate system.

If the light travels to another place, Δx^0 will remain constant. This Δx^0 will not be the same as the period of the same spectral line emitted by a local atom, which would be Δs again. The period is thus dependent on the gravitational potential g_{00} at the place where the light was emitted:

$$\Delta x^0 :: {g_{00}}^{-1/2}.$$

The spectral line will be shifted by this factor ${g_{00}}^{-1/2}$

If we use the Newtonian approximation (16.6), we have

$$\Delta x^0 :: 1 - V.$$

V will be negative at a place with a strong gravitational field, such as the surface of the sun, so, light emitted there will be red-shifted when compared with the corresponding light emitted on earth. The effect can be observed with the sun's light but it is rather masked by other physical effects, such as the Doppler effect arising from the motion of the emitting atoms. It can be better observed in light emitted from a white dwarf star, where the high density of the matter in the star gives rise to a much stronger gravitational potential at its surface.([11])

18 The Schwarzschild solution

The Einstein equations for empty space are nonlinear and are therefore very complicated, and it is difficult to get accurate solutions of them. There is, however, a special case which can be solved without too much trouble; namely the static spherically symmetric field produced by a spherically symmetric body at rest.

The static condition means that, with a static coordinate system, the $g_{\mu\nu}$ are independent of the time x^0 or t and also g_{0m}([12]). The spatial coordinates

[11] At the surface of a white dwarf star, precisely because of the high density of the stellar matter, the *small curvature* approximation—and the concept of gravitational potential energy—are not applicable.

[12] Making the assumption that all intervals of the form $(\delta x^0, 0, 0, 0)$ are temporal and all interval of the form $(0, \delta x^1, \delta x^2, \delta x^3)$ are spatial.

may be taken to be spherical polar coordinates $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$. The most general form of ds^2 compatible with spherical symmetry is

$$ds^2 = U dt^2 - V dr^2 - W r^2 (d\theta^2 + \sin^2\theta d\phi^2),$$

where U, V, and W are functions of r only. We may replace r by any function of r without disturbing the spherical symmetry. We use this freedom to simplify things as much as possible, and the most convenient arrangement is to have $W = 1$. The expression for ds^2 may then be written as

$$ds^2 = e^{2\nu} dt^2 - e^{2\lambda} dr^2 - r^2 d\theta^2 - r^2 \sin^2\theta d\phi^2, \tag{18.1}$$

with ν and λ functions of r only. They must be chosen to satisfy the Einstein equations. We can read off the values of the $g_{\mu\nu}$ from (18.1), namely,

$$g_0 0 = e^{2\nu}, \quad g_{11} = e^{-2\lambda}, \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2\theta,$$

and

$$g_{\mu\nu} = 0 \text{ for } \mu \neq \nu.$$

We find

$$g^{00} = e^{-2\nu}, \quad g_{11} = e^{-2\lambda}, \quad g^{22} = -r^2, \quad g^{33} = -r^{-2} \sin^{-2}\theta,$$

and

$$g^{\mu\nu} = 0 \text{ for } \mu \neq \nu.$$

It is now necessary to calculate all the Christoffel symbols $\Gamma^\sigma_{\mu\nu}$. Many of them vanish. The ones that do not are([13]):

$$\begin{array}{ll}
\Gamma^1_{00} = \nu' e^{2\nu - 2\lambda} & \Gamma^0_{10} = \nu' \\
\Gamma^1_{11} = \lambda' & \Gamma^2_{12} = \Gamma^3_{13} = r^{-1} \\
\Gamma^1_{22} = -r e^{-2\lambda} & \Gamma^2_{23} = \cot\theta \\
\Gamma^1_{33} = -r \sin^2\theta e^{-2\lambda} & \Gamma^2_{33} = -\sin\theta\cos\theta
\end{array}$$

These expressions are to be substituted in (14.4). The results are

$$R_{00} = \left(-\nu'' + \lambda'\nu' - \nu'^2 - \frac{2\nu'}{r} \right) e^{2\nu - 2\lambda}, \tag{18.2}$$

[13]With primes denoting differentiation with respect to r.

$$R_{11} = \nu'' - \lambda'\nu' + \nu'^2 - \frac{2\lambda'}{r}, \tag{18.3}$$

$$R_{22} = (1 + r\nu' - r\lambda')e^{-2\lambda} - 1, \tag{18.4}$$

$$R_{33} = R_{22} \sin^2 \theta,$$

with the other components of $R_{\mu\nu}$ vanishing.

Einstein's law of gravitation requires these expressions to vanish. The vanishing of (18.2) and (18.3) leads to

$$\lambda' + \nu' = 0.$$

For large values of r the space must approximate to being flat, so that λ and ν both tend to zero as $r \to \infty$. It follows that

$$\lambda + \nu = 0.$$

The vanishing of (18.4) now gives

$$(1 + 2r\nu')e^{2\nu} = 1$$

or

$$\left(re^{2\nu}\right)' = 1.$$

Thus

$$re^{2\nu} = r - 2m,$$

where m is a constant of integration. This also makes (18.2) and (18.3) vanish. We now get

$$g_{00} = 1 - \frac{2m}{r} \tag{18.5}$$

The Newtonian approximation must hold for large values of r. Comparing (18.5) with 16.6), we see that the constant of integration that has appeared in (18.5) is just the mass of the central body which is producing the gravitational field.

Te complete solution is

$$ds^2 = \left(1 - \frac{2m}{r}\right) - \left(1 - \frac{2m}{r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\phi^2. \tag{18.6}$$

It is known as the Schwarzschild solution. It holds outside the surface of the body that is producing the field, where there is no matter. Thus it holds fairly accurately outside the surface of a star.

The solution (18.6) leads to small corrections in the Newtonian theory for the motions of the planets around the Sun. These corrections are appreciable only in the case of Mercury, the nearest planet, and they explain the discrepancy of the motion of this planet with the Newtonian theory. Thus they provide a striking confirmation of the Einstein theory.[14]

19 Black holes

The solution (18.6) becomes singular at $r = 2m$, because then $g_{00} = 0$ and $g_{11} = -\infty$. It would seem that $r = 2m$ gives a minimum radius for a body of mass m. But a close investigation shows that this is not so.

Consider a particle falling into the central body and let its velocity vector be $v^\mu = dz^\mu/ds$. Let us suppose that it falls in radially, so that $v^2 = v^3 = 0$. The motion is determined by the geodesic equation (8.3):

$$\begin{aligned} \tfrac{dv^0}{ds} &= -\Gamma^0_{\mu\mu} v^\mu v^\nu = -g^{00}\Gamma_{0\mu\nu} v^\mu v^\nu \\ &= -g^{00} g_{00,1} v^0 v^1 = -g^{00} \tfrac{dg_{00}}{ds} v^0. \end{aligned}$$

Now $g^{00} = \frac{1}{g_{00}}$, so we get

$$g_{00}\frac{dv^0}{ds} + \frac{dg_{00}}{ds} v^0 = 0.$$

This integrates to

$$g_{00} v^0 = k,$$

with k a constant. It is the value of g_{00} where the particle starts to fall.

Again, we have

$$1 = g_{\mu\nu} v^\mu v^\nu = g_{00} {v^0}^2 + g_{11} {v^1}^2.$$

Multiplying this equation by g_{00} and using $g_{00} g_{11} = -1$, which we obtained in the last section, we find

$$k^2 - {v^1}^2 = g_{00} = 1 - \frac{2m}{r}.$$

[14]As we know that the mass of the Sun is not constant and this fact must have its effect on the displacement of the perihelion of the orbit of Mercury. I have attempted to compute this but the methods of classical perturbation theory cannot be applied because, when we consider a non constant solar mass, the resulting dynamical system stops being autonomous. Also, the displacement of the orbit of Mercury is small 39″ per century, i.e. 0.094″ per revolution (there are 1,296,000 seconds of arc in a complete revolution).

For a falling body $v^1 < 0$ and hence

$$v^1 = -\left(k^2 - 1 + \frac{2m}{r}\right)^{\frac{1}{2}}.$$

Now

$$\frac{dt}{dr} = \frac{v^0}{v^1} = -k\left(1 - \frac{2m}{r}\right)^{-1}\left(k^2 - 1 + \frac{2m}{r}\right)^{-\frac{1}{2}}.$$

Let us suppose the particle is close to the critical radius, so $r = 2m + \epsilon$ with ϵ small, and let us neglect ϵ^2. Then

$$\frac{dt}{dr} \approx -\frac{2m}{\epsilon} = -\frac{2m}{r - 2m}.$$

This integrates to

$$t = -2m\log(r - 2m) + \text{ constant}.$$

Thus, as $r \to 2m$ $t \to \infty$. The particle takes an infinite time to reach the critical radius $r = 2m$.

Let us suppose the particle is emitting light of a certain spectral line, and is being observed by someone at a large value of r. The light is red-shifted by a factor of ${g_{00}}^{-1/2}$. This factor becomes infinite as the particle approaches the critical radius. All physical processes on the particle will be observed to be going more and more slowly as it approaches $r = 2m$.

Now consider an observer traveling with the particle. His time scale is measured by ds. Now

$$\frac{ds}{dr} = \frac{1}{v^1} = -\left(k^2 - 1 + \frac{2m}{r}\right)^{-1/2},$$

and this tends to $-k^{-1}$ as r tends to $2m$. Thus, the particle reaches $r = 2m$ after a lapse of finite proper time for the observer. The traveling observer has aged only a finite amount when he reaches $r = 2m$. What will happen to him afterwards? He may continue sailing through empty space into smaller values of r.

To examine the continuation of the Schwarzschild solution for values of $r < 2m$, it is necessary to use a non static system of coordinates, so that we have the $g_{\mu\nu}$ varying with the time coordinate. We keep the coordinates θ and ϕ unchanged, but instead of t and r we use τ and ρ, defined by

$$\tau = t + f(r)\,\rho = t + g(r), \tag{19.1}$$

where the functions f and g are at our disposal.

We have, using the prime again to denote derivative with respect to r,

$$\begin{aligned} d\tau^2 - \tfrac{2m}{3}d\rho^2 &= (dt + f'dr)^2 - \tfrac{2m}{r}(dt + g'dr)^2 \\ &= \left(1 - \tfrac{2m}{r}\right)dt^2 + 2\left(f' - \tfrac{2m}{r}g'\right)dtdr + \left(f'^2 - \tfrac{2m}{r}g'^2\right)dr^2 \\ &= \left(1 - \tfrac{2m}{r}\right)dt^2 - \left(1 - \tfrac{2m}{r}\right)^{-1}dr^2 \end{aligned} \tag{19.2}$$

provided we choose the functions f and g to satisfy

$$f' = \frac{2m}{r}g' \tag{19.3}$$

and

$$\frac{2m}{r}g'^2 - f'^2 = \left(1 - \frac{2m}{r}\right)^{-1}. \tag{19.4}$$

Elimination of f from these equations gives

$$g' = \left(\frac{r}{2m}\right)^{1/2}\left(1 = \frac{2m}{r}\right)^{-1} \tag{19.5}$$

To integrate this equation, put $r = y^2$ and $2m = a^2$. With $r > 2m$ we have $y > a$. We now have

$$\frac{dg}{dy} = 2y\frac{dg}{dr} = \frac{2y^4}{a}\frac{1}{y^2 - a^2},$$

$$g = \frac{2}{3a}y^3 + 2ay - a^2\log\frac{y+1}{y-a}. \tag{19.6}$$

Finally we get from (19.3) and (19.5)

$$g' - f' = \left(1 - \frac{2m}{r}\right)g' = \left(\frac{r}{2m}\right)^{1/2},$$

which integrates to

$$\frac{2}{3}\frac{1}{\sqrt{2m}}r^{3/2} = g - f = \rho - \tau. \tag{19.7}$$

Thus

$$r = \mu(\rho - \tau)^{2/3}, \tag{19.8}$$

with

$$\mu = \left(\frac{3}{2}\sqrt{2m}\right)^{2/3}.$$

In this way we see that we can satisfy the conditions (19.3) and (19.4) and so we can use (19.2). Substituting into the Schwarzschild solution (18.6), we get

$$ds^2 = d\tau^2 - \frac{2m}{\mu(\rho-\tau)^{2/3}}d\rho^2 - \mu^2(\rho-\tau)^{4/3}\left(d\theta^2 + \sin^2\theta d\phi^2\right). \tag{19.9}$$

The critical value $r = 2m$ corresponds, from (19.7), to $\rho - \tau = 4m/3$. There is no singularity here in the metric (19.9).

We know that the metric (19.9) satisfies the Einstein equations for empty space in the region $r > 2m$, because it can be transformed to the Schwarzschild solution by a mere change of coordinates. We can infer that it satisfies the Einstein equations for $r \leq 2m$ from analytic continuity, because it does not involve any singularity at $r = 2m$. It may continue to hold down to $r = 0$ or $\rho - \tau = 0$. The singularity appears in the connection between the new coordinates and the original ones, equation (19.1). But once we have established the coordinate system we can disregard the previous one and the singularity no longer appears.

We see that the Schwarzschild solution for empty space can be extended to the region $r < 2m$. But this region cannot communicate with the space for which $r > m$. Any signal, even a light signal, will take an infinite time to reach the boundary $r = 2m$, as we can easily check. Thus we cannot have direct observational knowledge of the region $r < 2m$. Such a region is called a black hole, because things can fall into it (taking and infinite time by our clocks, to do so) but nothing can come out.

The question arises whether such a region can actually exist. All we can say is that definitely the Einstein equations allow it. A massive stellar object can collapse to a very small radius and the gravitational forces then become so strong that no physical forces can hold them in check and prevent further collapse. It would seem that it would have to collapse into a black hole. It would take an infinite time to do so by our clocks, but only a finite time relative to the collapsing matter itself.

20 Tensor densities

With a transformation of coordinates, an element of four-dimensional volume transforms according to the law

$$dx^{0'}dx^{1'}dx^{2'}dx^{3'} = dx^0 dx^1 dx^2 dx^3 J, \tag{20.1}$$

where J is the Jacobian

$$J = \frac{\partial(x^{0'}, x^{1'}, x^{2'}, x^{3'})}{\partial(x^0, x^1, x^2, x^3)} = \text{ determinant of } x^{\mu'}_{,\nu}$$

We may write (20.1)

$$d^4x' = J d^4x \tag{20.2}$$

for brevity.

Now

$$g_{\alpha\beta} = x^{\mu'}_{,\alpha} g_{\mu'\nu'} x^{\nu'}_{,\beta}.$$

We can look upon the right side as the product of three matrices, the first matrix having its rows specified by α and columns specified by μ', the second having its rows specified by μ' and columns by ν' and the third having its rows specified by ν' and columns by β. This product equals the matrix $g_{\alpha\beta}$ on the left. The corresponding equation must hold between the determinants; therefore

$$g = J g' J$$

or

$$g = J^2 g'.$$

Now, g is a negative quantity, so we may form $\sqrt{-g}$, taking the positive value for he square root. Thus

$$\sqrt{-g} = J\sqrt{-g'}. \tag{20.3}$$

Suppose S is a scalar field quantity, $S = S'$. Then

$$\int S\sqrt{-g}d^4x = \int S\sqrt{-g'}Jd^4x = \int S'\sqrt{-g'}d^4x',$$

if the region of integration for the x' corresponds to that for the x. Thus

$$\int S\sqrt{-g}d^4x = \text{ invariant.} \tag{20.4}$$

We call $S\sqrt{-g}$ a scalar density, meaning a quantity whose integral is invariant.

Similarly, for any tensor field $T^{\mu\nu\cdots}$ we may call $T^{\mu\nu\cdots}\sqrt{-g}$ a tensor density. The integral

$$\int T^{\mu\nu\cdots}\sqrt{-g}d^4x$$

is a tensor if the domain of integration is small. It is not a tensor if the domain of integration is not small, because it then consists of a sum of tensors located at different points and it does not transform in any simple way in a transformation of coordinates.

The quantity $\sqrt{-g}$ will be very much used in the future. For brevity we shall write it simply as $\surd$. We have ([15])

$$g^{-1}g_{,\nu} = 2\surd^{-1}\surd_{,\nu},$$

Thus the formula (14.5) gives

$$\surd_{,\nu} = \frac{1}{2}\surd g^{\lambda\mu}g_{\lambda\mu,\nu} \tag{20.5}$$

and the formula (14.6) may be written

$$\Gamma^{\mu}_{\nu\mu}\surd = \surd_{,\nu} \tag{20.6}$$

21 Gauss and Stokes theorems

The vector A^μ has the covariant divergence $A^\mu_{;\mu}$, which is a scalar. We have

$$A^{\mu}{}_{;\mu} = A^\mu_{,\mu} + \Gamma^\mu_{\nu\mu}A^\nu = A^\mu_{,\mu} + \surd^{-1}\surd_{,\nu}A^\nu.$$

Thus

$$A^\mu_{;\mu}\surd = \left(A^\mu\surd\right)_{,\mu}. \tag{21.1}$$

We can put $A^\mu_{;\mu}\surd$ for S in (20.4), and we get the invariant

$$\int A^\mu_{;\mu}\surd d^4x = \int \left(A^\mu\surd\right)_{,\mu} d^4x.$$

[15] This follows from the identity $\log\sqrt{-g} = \frac{1}{2}\log(-g)$.

If the integral is taken over a finite (four-dimensional) volume, the right hand side can be converted by Gauss's theorem to an integral over the boundary surface (three-dimensional) of that volume.

If $A^{\mu}_{;\mu} = 0$, we have

$$\left(A^{\mu}\sqrt{}\right)_{,\mu} = 0 \tag{21.2}$$

and this gives us a conservation law; namely, the conservation of a fluid whose density is $A^0\sqrt{}$ and whose flow is given by the three-dimensional vector $A^m\sqrt{}(m = 1, 2, 3)$. We may integrate (21.2) over a three-dimensional volume V lying at a fixed time x^0([16]). The result is

$$\begin{aligned}\left(\int A^0\sqrt{}d^3x\right)_{,0} &= -\int (A^m)_{,m}\, d^3x\\ &= \text{surface integral over boundary of } V\end{aligned}$$

If there is no current in the boundary of V, $\int A^0\sqrt{}d^3x$ is constant.

Those results for a vector A^{μ} cannot be taken over to a a tensor with more than on suffix, in general. Take a two suffix tensor $Y^{\mu\nu}$. In flat space we can use Gauss's theorem to express $\int Y^{\mu\nu}{}_{,\nu}dx^4$ as a surface integral, but in curved space we cannot in general express $\int Y^{\mu\nu}{}_{;\nu}\sqrt{}dx^4$ as a surface integral. An exception occurs for an antisymmetrical tensor $F^{\mu\nu} = -F^{\nu\mu}$.

In this case we have

$$F^{\mu\nu}{}_{;\sigma} = F^{\mu\nu}{}_{,\sigma} + \Gamma^{\mu}_{\sigma\rho}F^{\rho\nu} + \Gamma^{\nu}_{\sigma\rho}F^{\mu\rho}$$

so

$$\begin{aligned}F^{\mu\nu}{}_{;\nu} &= F^{\mu\nu}{}_{,\nu} + \Gamma^{\mu}_{\nu\rho}F^{\rho\nu} + \Gamma^{\nu}_{\nu\rho}F^{\mu\rho}\\ &= F^{\mu\nu}{}_{,\nu} + \sqrt{}^{-1}\sqrt{}_{,\rho}F^{\mu\rho}\end{aligned}$$

from (20.6) ([17]). Thus

$$F^{\mu\nu}{}_{;\nu}\sqrt{} = \left(F^{\mu\nu}\sqrt{}\right)_{,\nu} \tag{21.3}$$

Hence $\int F^{\mu\nu}{}_{;\nu}\sqrt{}d^4x =$ a surface integral, and if $F^{\mu\nu}{}_{;\nu} = 0$ we have a conservation law.

In the symmetrical case $Y^{\mu\nu} = Y^{\nu\mu}$ we can get a corresponding equation with an extra term, provided we put one of the suffixes downstairs and deal with $Y_{\mu}{}^{\nu}{}_{;\nu}$. We have

$$Y_{\mu}{}^{\nu}{}_{;\sigma} = Y_{\mu}{}^{\nu}{}_{,\sigma} - \Gamma^{\alpha}_{\mu\sigma}Y_{\alpha}{}^{\nu} + \Gamma^{\nu}_{\sigma\alpha}Y_{\mu}{}^{\alpha}.$$

[16] I sense here an error of method because Dirac makes the assumption that x^0 is strictly temporal.

[17] And the fact that $\Gamma^{\mu}_{\nu\rho}F^{\rho\nu} \equiv 0$ because $\Gamma^{\mu}_{\nu\rho} = \Gamma^{\mu}_{\rho\nu}$.

Putting $\sigma = \nu$ and using (20.6), we get

$$Y_{\mu}{}^{\nu}{}_{;\nu} = Y_{\mu}{}^{\nu}{}_{,\nu} + \sqrt{}^{(-1)}\sqrt{}_{\alpha} Y_{\mu}{}^{\alpha} - \Gamma_{\alpha\mu\nu} Y^{\alpha\nu}.$$

Since $Y^{\mu\nu}$ is symmetrical, we can replace the $\Gamma_{\alpha\mu\nu}$ in the last term by

$$\frac{1}{2}\left(\Gamma_{\alpha\nu\mu} + \Gamma_{\nu\alpha\mu}\right) = \frac{1}{2} g_{\alpha\nu,\mu}$$

from (7.6). Thus we get

$$Y_{\mu}{}^{\nu}{}_{;\nu} = Y_{\mu}{}^{\nu}{}_{,\nu} + \left(Y_{\mu}{}^{\nu}\sqrt{}\right)_{,\nu} - \frac{1}{2} g_{\alpha\beta,\mu} Y^{\alpha\beta}. \tag{21.4}$$

For a covariant vector A_μ, we have

$$\begin{aligned} A_{\mu;\nu} - A_{\nu;\mu} &= A_{\mu,\nu} - \Gamma^{\rho}_{\mu\nu} A_\rho - \left(A_{\nu,\mu} - \Gamma^{\rho}_{\nu\mu} A_\rho\right) \\ &= A_{\mu,\nu} - A_{\nu,\mu}. \end{aligned} \tag{21.5}$$

Let us integrate this equation over an area of the surface $x^0 =$ constant, $x^3 =$ constant. From Stokes theorem we get

$$\begin{aligned} \iint \left(A_{1;2} - A_{2;1}\right) dx^1 dx^2 &= \iint \left(A_{1,2} - A_{2,1}\right) dx^1 dx^2 \\ &= \oint \left(A_1 dx^1 + A_2 dx^2\right) \end{aligned} \tag{21.6}$$

integrated around the perimeter of the area. Thus we get an integral round a perimeter equated to a flux crossing the surface bounded by the perimeter.

The result must hold in all coordinate systems,not merely those for which the equations of the surface are $x^0 =$ constant, $x^3 =$ constant.

To get an invariant way of writing the result, we introduce a general formula for an element of two-dimensional surface. If we take two small contravariant vectors ξ^μ and ζ^ν, the element of surface area that they subtend is determined by the antisymmetric two-index tensor

$$dS^{\mu\nu} = \xi^\mu \zeta^\nu - \xi^\nu \zeta^\mu,$$

Thus, if ξ^μ has the components $0, dx^1, 0, 0$ and ζ^ν has the components $0, 0, dx^2, 0$, then $dS^{\mu\nu}$ has the components

$$dS^{12} = dx^1 dx^2,\ dS^{21} = -dx^1 dx^2,$$

with the other components vanishing. The left-hand side of (21.6) becomes

$$\iint A_{\mu;\nu} dS^{\mu;\nu}.$$

The right hand side is evidently $\oint A_\mu dx^\mu$, so the formula becomes ([18])

$$\frac{1}{2}\iint \left(A_{\mu;\nu} - A_{\nu;\mu}\right) dS^{\mu\nu} = \oint A_\mu dx^\mu. \tag{21.7}$$

22 Harmonic coordinates

The d'Alembert equation for a scalar V, namely $\Box V = 0$, gives, from (10.8),

$$g^{\mu\nu}\left(V_{\mu\nu} - \Gamma^\alpha_{\mu\nu} V_{,\alpha} = 0.\right) \tag{22.1}$$

If we are using rectilinear axes in flat space, each of the four coordinates x^λ satisfies $\Box x^\lambda = 0$. We might substitute x^λ for V in (22.1). The result, of course, is not a tensor equation, because x^λ is not a scalar like V, so it holds only in certain coordinate systems. It imposes a restriction on the coordinates.

If we substitute x^λ for V, then for $V_{,\alpha}$ we must substitute $x^\lambda_{,\alpha} = g^\lambda_\alpha$([19]). The equation (22.1) becomes

$$g^{\mu\nu}\Gamma^\lambda_{\mu\nu} = 0. \tag{22.2}$$

Coordinates that satisfy this condition are called *harmonic coordinates*. They provide the closest approximation to rectilinear coordinates that we can have in curve space. They provide the closes approximation to rectilinear coordinates that we can have in curved space. We may use them in any problem if we wish to, but very often they are not worthwhile because the tensor formalism with general coordinates is really quite convenient. For the discussion of gravitational waves, however, harmonic coordinates are very useful.

We have in general coordinates, from (7.9) and (7.6),

$$\begin{aligned} {g^{\mu\nu}}_{,\sigma} &= -g^{\mu\alpha}g^{\nu\beta}\left(\Gamma_{\alpha\beta\sigma} + \Gamma_{\beta\alpha\sigma}\right) \\ &= -g^{\nu\beta}\Gamma^\mu_{\beta\sigma} - g^{\mu\alpha}\Gamma^\nu_{\alpha\sigma}. \end{aligned} \tag{22.3}$$

[18] The first integral on the surface and the second on the perimeter.

[19] Which is a constant and therefore $x^\lambda_{,\alpha\beta} \equiv 0$.

Thus, with the help of (20.6),

$$\left(g^{\mu\nu}\surd\right)_{,\sigma} = \left(-g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} - g^{\mu\alpha}\Gamma^{\nu}_{\alpha\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta}\right)\surd. \tag{22.4}$$

Contracting by putting $\sigma = \nu$, we get

$$\left(g^{\mu\nu}\surd\right)_{,\nu} = -g^{\nu\beta}\Gamma^{\mu}_{\beta\nu}\surd. \tag{22.5}$$

We see now that an alternative form for the harmonic condition is

$$\left(-g^{\mu\nu}\surd\right)_{,\nu} = 0. \tag{22.6}$$