21 October 2011

General Relativity: An Informal Primer

David Kaiser

Center for Theoretical Physics, MIT

1 Introduction

General relativity, and its application to cosmological models such as inflation, is a remarkably beautiful and elegant theory. Yet newcomers to the field often face at least three types of challenges: conceptual, mathematical, and notational. Over the past century, physicists and mathematicians have developed a remarkably compact formalism and notation for working with general relativity, which requires some practice to really get comfortable with. The methods are quite powerful and well worth the investment of effort. One can only really hope to tackle the fascinating conceptual and mathematical challenges if one has first gotten over the hurdle of understanding how to parse and manipulate the formalism.

These notes are intended to give an informal and very brief introduction to working with general relativity. They are certainly no substitute for a real textbook; rather, they are geared to be read in conjunction with a book like Steven Weinberg's *Gravitation and Cosmology* [1], to provide a little extra practice unpacking the unfamiliar notation so that students may dig in to the real meat of Weinberg's book. Think of these notes, in other words, like "language lab": before one can engage at a deep level with the novels of Molière or the philosophy of Hegel, one must gain at least a rudimentary familiarity with the medium of composition.

Weinberg's textbook is aimed at graduate students. Some more accessible books, aimed at advanced undergraduates, have appeared in recent years, and these might be helpful as additional references, including Wolfgang Rindler's *Relativity: Special, General, and Cosmological* [2] and Tai-Pei Cheng's *Relativity, Gravitation, and Cosmology: A Basic Introduction* [3]. Cheng's notation is closer to Weinberg's, and hence to the notation we will be using. Rindler introduces his own notation that has some advantages on its own terms but which does not exactly match that of Weinberg or the others.

2 Special Relativity Basics

Imagine how cumbersome it would be to study electricity and magnetism without the use of 3-vectors. Rather than write compact terms such as $\mathbf{E}$ or $\mathbf{B}$ for the electric and magnetic fields, one would always have to work in terms of the fields' components within a given coordinate system, such as E_x or B_z. In fact, this is precisely how Maxwell himself manipulated his own equations in the 1850s and 1860s, and a large part of the reason why his famous *Treatise on Electricity and Magnetism*, published in 1873, filled two hulking volumes and totaled nearly 1000 pages [4]. The vector notation we all enjoy today was invented by the British physicist and engineer Oliver Heaviside in the 1880s expressly to simplify manipulations of Maxwell's equations [5]. If Heaviside's vector notation had been available to Maxwell when he was writing his *Treatise*, the entire tome could have been expressed, if not in a single tweet, then certainly within a few modest blog posts.

A similar dramatic streamlining comes from introducing relativistic 4-vectors and tensors, which will become crucial for manipulating quantities of interest within special and general relativity. Like Heaviside's 3-vector notation, the relativistic notation can both save space on a page as well as clarify relationships between different physical concepts or quantities. But just like the 3-vector notation, the compactness achieved by using relativistic notation takes some practice getting used to. To really appreciate the power and efficiency of the notation, one must practice unpacking the relationships.

2.1 4-vectors

One of the main consequences of Einstein's special relativity is that we do not view space and time as inherently distinct from each other. To take advantage of a genuinely four-dimensional view, we use 4-vectors in place of Heaviside's 3-vectors, which had been restricted to space-like quantities. For example, one can consider a spatial vector $\mathbf{x} = (x, y, z)$, which we can write as $x^i = (x, y, z)$, where the index, i, labels the individual components. The index i ranges between 1 and 3, since 3-vectors contain just three components. In this example, $x^1 = x$, $x^2 = y$, and $x^3 = z$.

We can construct similar objects that include both time-like and spatial components. To distinguish these 4-vectors from the usual 3-vectors, we will use Greek letters as indices. Thus if we continue to label the spatial coordinates in Cartesian form, via $x^i = (x, y, z)$, then our spacetime 4-vector becomes

$$x^\mu = (ct, x^i) = (ct, x, y, z), \tag{2.1}$$

where c is the speed of light. By convention, we will label the first component of our 4-vectors as the "0th" component. Thus Greek-letter indices range over $\mu, \nu = 0, 1, 2, 3$, while lowercase Latin indices range over $i, j = 1, 2, 3$. In our present example, $x^0 = ct$. Four-vectors with their index in the upper position, as in Eq. (2.1), are known as "contravariant" vectors.

Next we may introduce a convenient array of numbers: not a 4-vector but a 4×4 matrix, via the relations

$$\eta_{\mu\nu} = \begin{cases} -1\,, & \mu = \nu = 0, \\ +1\,, & \mu = \nu = 1,\ 2,\ \text{or } 3, \\ 0\,, & \mu \neq \nu. \end{cases} \tag{2.2}$$

We may write out $\eta_{\mu\nu}$ in explicit matrix form,

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{2.3}$$

though usually the abstract representation of Eq. (2.2) will prove most convenient.

We may also construct the inverse of this matrix, $\eta^{\mu\nu}$, with its indices in the upper position rather than the lower position. The matrix $\eta^{\mu\nu}$ must satisfy the relations

$$\eta^{\mu\nu}\eta_{\mu\alpha} = \delta^{\nu}{}_{\alpha}. \tag{2.4}$$

The symbol on the righthand side is just the usual Kronecker delta function, which satisfies $\delta^{\nu}{}_{\alpha} = 1$ if $\nu = \alpha$, and $\delta^{\nu}{}_{\alpha} = 0$ if $\nu \neq \alpha$.

Eq. (2.4) is our first instance of the special compactness of the relativistic notation. It is written with what is known as the "Einstein summation convention": whenever an index appears in both an upper and a lower position, we must sum over that repeated index. Thus the expression on the lefthand side of Eq. (2.4) is actually a shorthand for the following sum of terms:

$$\eta^{\mu\nu}\eta_{\mu\alpha} = \eta^{0\nu}\eta_{0\alpha} + \eta^{1\nu}\eta_{1\alpha} + \eta^{2\nu}\eta_{2\alpha} + \eta^{3\nu}\eta_{3\alpha}. \tag{2.5}$$

Using the definition of $\eta_{\mu\nu}$ in Eq. (2.2), we now see that the first term in the sum, proportional to $\eta_{0\alpha}$, will vanish for all values of the index α except for $\alpha = 0$; likewise, the second term in the sum will vanish for all values of the index α except for $\alpha = 1$, and so on. We therefore find, for example,

$$\eta^{0\nu} = \frac{1}{(-1)}\delta^{\nu}{}_{0}. \tag{2.6}$$

Proceeding similarly for each value of ν and α, Eq. (2.4) implies that

$$\eta^{\mu\nu} = \begin{cases} -1\,, & \mu = \nu = 0, \\ +1\,, & \mu = \nu = 1,\ 2,\ \text{or } 3, \\ 0\,, & \mu \neq \nu. \end{cases} \tag{2.7}$$

We may use these matrices, $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$, to raise and lower indices on other objects, such as 4-vectors. For example, from Eq. (2.1) we may construct a 4-vector with its index in the lower position, which is known as a "covariant vector":

$$x_\mu \equiv \eta_{\mu\nu}\, x^\nu. \tag{2.8}$$

Again recalling the Einstein summation convention, Eq. (2.8) is actually a shorthand for the sum of terms,

$$\begin{aligned} x_\mu &= \eta_{\mu 0} x^0 + \eta_{\mu 1} x^1 + \eta_{\mu 2} x^2 + \eta_{\mu 3} x^3 \\ &= (-ct, x, y, z), \end{aligned} \tag{2.9}$$

where the second line follows from the form of $\eta_{\mu\nu}$ in Eq. (2.2), since $\eta_{\mu\nu}$ only has nonzero components along the diagonal, for $\mu = \nu$.

Indices that are repeated, and hence summed over, are often called "dummy" indices. One can change the letter used for those repeated indices without making any change to the actual content of the expressions. For example,

$$\eta_{\mu\nu}\, x^\nu = \eta_{\mu\alpha}\, x^\alpha, \tag{2.10}$$

since in each case, the repeated index is summed over, so the final answer is independent of that index. One can use the notation to perform a quick spot-check when in the midst of lengthy equations: do the "free" indices — that is, the indices that are *not* repeated within a given expression, and hence not summed over — agree on both sides of the equation? In Eq. (2.10), the expression on the lefthand side has a single "free" index, μ, in a lower position, and no "free" indices in an upper position; and so does the expression on the righthand side. Getting the "free" indices to match on each side is a bare minimum in making sure a given equation is correct.

We may also generalize the dot product of two 3-vectors to the case of 4-vectors, by using the matrix $\eta_{\mu\nu}$:

$$\begin{aligned} x^2 &= x_\mu x^\mu = \eta_{\mu\nu}\, x^\mu x^\nu \\ &= \eta_{00} x^0 x^0 + \eta_{01} x^0 x^1 + \eta_{02} x^0 x^2 + \eta_{03} x^0 x^3 \\ &\qquad + \eta_{10} x^1 x^0 + \eta_{11} x^1 x^2 + \eta_{12} x^1 x^2 + \eta_{13} x^1 x^3 \\ &\qquad + \eta_{20} x^2 x^0 + \eta_{21} x^2 x^1 + \eta_{22} x^2 x^2 + \eta_{23} x^2 x^3 \\ &\qquad + \eta_{30} x^3 x^0 + \eta_{31} x^3 x^1 + \eta_{32} x^3 x^2 + \eta_{33} x^3 x^3 \\ &= -c^2 t^2 + x^2 + y^2 + z^2. \end{aligned} \tag{2.11}$$

The middle four lines come from applying the Einstein summation convention: every repeated index gets summed over. The final line follows from the particular form of $\eta_{\mu\nu}$, and from the vector x^μ in Eq. (2.1). It should be clear from context that the "x^2" that appears on the lefthand side of the top line refers to the square of the 4-vector named x^μ, whereas the "x^2" that appears on the righthand side in the middle rows refers to the $\mu = 2$ component of the vector x^μ. In general it can take a little practice to get used to distinguishing algebraic symbols, like squares and cubes, from component indices on contravariant vectors.

Another general feature is that we can "see-saw" pairs of repeated indices, swapping upper and lower indices without changing the result. For any general vectors A^μ and B^ν, their dot product is unchanged whether we write $A^\mu B_\mu$ or $A_\mu B^\mu$, since

$$A^\mu B_\mu = \eta_{\mu\nu} A^\mu B^\nu = A_\nu B^\nu. \tag{2.12}$$

We also see that the Kronecker delta function multiplying an expression effectively replaces a summed-over index with the other index attached to the delta function. For an arbitrary 4-vector A^μ,

$$\delta^\nu{}_\mu A^\mu = A^\nu. \tag{2.13}$$

The same holds for more complicated expressions, such as

$$\delta^\mu{}_\nu \delta^\lambda{}_\rho T^{\alpha\nu} B_{\beta\sigma} C^\rho = T^{\alpha\mu} B_{\beta\sigma} C^\lambda. \tag{2.14}$$

As usual, these expressions can be spot-checked using the rule of matching free indices on each side.

2.2 Coordinate transformations

Consider two reference frames, S and S', moving relative to each other at a constant speed v along the $\hat{x}$ direction. In students' first encounter with special relativity, they typically see the coordinate transformation between frames S and S' written as

$$\begin{aligned} t \to t' &= \gamma \left(t - \frac{vx}{c^2} \right), \\ x \to x' &= \gamma \left(x - vt \right), \\ y \to y' &= y, \\ z \to z' &= z. \end{aligned} \tag{2.15}$$

For now, we will consider γ some function of the relative velocity between the two frames, v, but otherwise leave its form unspecified. We will also scale our units for space and time such that the speed of light is set to 1: $c = 1$.

We may write the Lorentz transformation of Eq. (2.15) in more compact form if we introduce the 4-vector of spacetime coordinates, x^μ, as in Eq. (2.1). The transformed coordinates correspond to a new vector, x'^{μ}, which may be related to the original vector via the relation

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}\, x^{\nu}. \tag{2.16}$$

Again it is worthwhile to unpack exactly what Eq. (2.16) stands for: it is a shorthand notation for a sum of terms,

$$\begin{aligned} x'^{\mu} &= \Lambda^{\mu}{}_{0}x^{0} + \Lambda^{\mu}{}_{1}x^{1} + \Lambda^{\mu}{}_{2}x^{2} + \Lambda^{\mu}{}_{3}x^{3} \\ &= \Lambda^{\mu}{}_{0}t + \Lambda^{\mu}{}_{1}x + \Lambda^{\mu}{}_{2}y + \Lambda^{\mu}{}_{3}z, \end{aligned} \tag{2.17}$$

where the second line follows for our original choice of coordinates, $x^{\mu} = (t, x, y, z)$.

For the transformation of Eq. (2.15), the individual components of $\Lambda^{\mu}{}_{\nu}$ may easily be solved for. The relation between t' and t in Eq. (2.15), for example, means that we have

$$\begin{aligned} \Lambda^{0}{}_{0} &= \gamma, \\ \Lambda^{0}{}_{1} &= -\gamma v, \\ \Lambda^{0}{}_{2} &= 0, \\ \Lambda^{0}{}_{3} &= 0. \end{aligned} \tag{2.18}$$

Proceeding similarly from the relations between x' and x, y' and y, and z' and z in Eq. (2.15), we may solve for all components of the array $\Lambda^{\mu}{}_{\nu}$:

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.19}$$

Now we may begin to appreciate the power of compact notation such as Eq. (2.16). The particular transformation of Eq. (2.15) was by no means the most general transformation we could have considered, even within special relativity. We could have made any combination of boost along a given direction as well as a rotation along any axis. Those more complicated transformations would have involved additional mixing among components of x'^{μ} and x^{μ}: a boost along an arbitrary direction combined with a rotation in the y-z plane, for example, would have meant that z' became a function of t, x, y, and z. But even that more complicated situation would still be represented in the compact form of Eq. (2.16); all that would change would be the particular values of the components of $\Lambda^{\mu}{}_{\nu}$. (See [1], section 2.1.)

Recall from special relativity the remarkable result that observers who are in motion with respect to each other — and who will therefore disagree about measurements of time intervals and spatial lengths — will nonetheless agree on a particular combination of time and space measurements. That combination is an invariant spacetime interval, and can be written,

$$s^2 = -\left(\Delta t\right)^2 + \left(\Delta \mathbf{x}\right)^2 = -\left(\Delta t\right)^2 + \left(\Delta x\right)^2 + \left(\Delta y\right)^2 + \left(\Delta z\right)^2. \tag{2.20}$$

(Remember that we are using units such that $c = 1$. Also note that Weinberg uses the letter τ for what I have labeled s, which is reasonable, since the quantity represents the "proper time.") Note the critical minus sign in front of the time piece. Conventions vary as to whether one puts the minus sign in front of $(\Delta t)^2$ or $(\Delta \mathbf{x})^2$; what matters most is the relative minus-sign between the two terms. A common choice in the literature on general relativity, and the choice I will use here, is to put the minus sign in front of the temporal term. Weinberg uses the opposite sign convention in [1]. Thus we must keep the sign conventions in mind when reading Weinberg's book.

The invariance of the interval s^2 means that s^2 is unchanged even if one makes a coordinate transformation, such as recalibrating one's clocks and meter sticks to those of a reference frame moving at some constant speed with respect to the observer. The invariance of the spacetime interval, s^2, restricts the forms that $\Lambda^\mu{}_\nu$ can take. For the transformation represented by Eq. (2.19), this means that the factor γ must take a particular form. Given the coordinate transformation of Eq. (2.15), we may consider each reference frame to share the same origin, that is, $t_0 = t'_0 = 0$, $x_0 = x'_0 = 0$, $y_0 = y'_0 = 0$, and $z_0 = z'_0 = 0$. Then the intervals in Eq. (2.20) take the form $(\Delta t)^2 = (t - t_0)^2 = t^2$, and hence

$$
\begin{aligned}
s'^2 &= -\left(\Delta t'\right)^2 + \left(\Delta \mathbf{x}'\right)^2 \\
&= -\gamma^2 \left(t - vx\right)^2 + \gamma^2 \left(x - vt\right)^2 + y^2 + z^2 \\
&= -\gamma^2 t^2 \left(1 - v^2\right) + \gamma^2 x^2 \left(1 - v^2\right) + y^2 + z^2 .
\end{aligned}
\tag{2.21}
$$

In order for s'^2 in Eq. (2.21) to match the invariant distance as calculated in the original frame, s^2 in Eq. (2.20), we require

$$
\gamma = \frac{1}{\sqrt{1 - v^2}}. \tag{2.22}
$$

For the simple transformation of Eq. (2.15), finding the required form of γ, and hence of $\Lambda^\mu{}_\nu$, was straightforward. But what if we had considered a more complicated coordinate transformation? One can appreciate how tedious it would become to repeat each of the steps, component by component.

Instead, upon using the matrix $\eta_{\mu\nu}$, we may rewrite the spacetime interval of Eq. (2.20) as

$$
\begin{aligned}
s^2 &= x_\mu x^\mu = \eta_{\mu\nu}\, x^\mu x^\nu \\
&= -t^2 + x^2 + y^2 + z^2 ,
\end{aligned}
\tag{2.23}
$$

where the second line follows from Eq. (2.11). We may similarly calculate the dot product of x'^μ:

$$
s'^2 = x'_\mu x'^\mu = \eta_{\mu\nu}\, x'^\mu x'^\nu = \eta_{\mu\nu} \left(\Lambda^\mu{}_\alpha x^\alpha\right) \left(\Lambda^\nu{}_\beta x^\beta\right) . \tag{2.24}
$$

The requirement that $s'^2 = s^2$ places a restriction on the form of $\Lambda^\mu{}_\nu$, namely,

$$\eta_{\mu\nu}\,\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = \eta_{\alpha\beta}. \tag{2.25}$$

So far we have been treating any list of four components as if it were a relativistic 4-vector, and we have not made any distinctions between arrays of numbers such as matrices and genuine tensors. In fact, not any list of four components is an actual 4-vector; and not every array that can be written as a matrix is an actual tensor. The key property that determines whether objects are vectors or tensors is how those objects behave under a coordinate transformation. (See [1], section 4.2.)

It is easiest to introduce the basic notion in the context of special relativity, for which the class of coordinate transformations under consideration is relatively simple, and many of the components of the relevant matrices are simply constants. Then we may build from these early examples to the case of arbitrary coordinate transformations, not restricted to the case of inertial reference frames (moving at constant speeds).

A genuine relativistic 4-vector transforms in a specific way under a coordinate transformation. Upon the transformation $x^\mu \to x'^\mu$, a contravariant 4-vector A^μ will transform to A'^μ, given by

$$A'^\mu = \left(\frac{\partial x'^\mu}{\partial x^\nu}\right) A^\nu, \tag{2.26}$$

whereas a covariant 4-vector A_μ will transform to A'_μ,

$$A'_\mu = \left(\frac{\partial x^\mu}{\partial x'^\nu}\right) A_\nu. \tag{2.27}$$

Remember as always that the Einstein summation convention means that the righthand sides of Eqs. (2.26) and (2.27) are actually sums of several terms each. That is, components of the transformed vector will become linear combinations of components of the original vector, just as t' and x' — components of the transformed 4-vector x'^μ — become linear combinations of components t and x of x^μ. For the specific coordinate transformation of Eq. (2.15), we have

$$\begin{aligned} A'^0 &= \left(\frac{\partial t'}{\partial x^\nu}\right) A^\nu \\ &= \left(\frac{\partial t'}{\partial t}\right) A^0 + \left(\frac{\partial t'}{\partial x}\right) A^1 + \left(\frac{\partial t'}{\partial y}\right) A^2 + \left(\frac{\partial t'}{\partial z}\right) A^3 \\ &= \gamma A^0 - \gamma v A^1, \end{aligned} \tag{2.28}$$

and similarly for the remaining components of A'^μ.

Tensors are generalizations of vectors. They have a "rank" given by the number of independent indices required to represent all of their components. For example, $B^{\mu\nu}$, $C^\mu{}_\nu$, and $D_{\mu\nu}$ would each represent rank-2 tensors; $E^{\mu\nu\sigma}$ is a rank-3 tensor. A tensor that has all of its indices in the upper position is known as a "contravariant tensor." Under a coordinate transformation, it transforms as

$$B'^{\mu\nu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\right)\left(\frac{\partial x'^{\nu}}{\partial x^{\beta}}\right) B^{\alpha\beta}. \tag{2.29}$$

Likewise, a tensor with all of its indices in the lower position is known as a "covariant tensor," and it transforms as

$$C'_{\mu\nu} = \left(\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\right)\left(\frac{\partial x^{\beta}}{\partial x'^{\nu}}\right) C_{\alpha\beta}. \tag{2.30}$$

Since we may always raise or lower individual indices using $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$, we may also work with mixed tensors, such as $B^{\mu}{}_{\nu}$, which transform as

$$B'^{\mu}{}_{\nu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\right)\left(\frac{\partial x^{\beta}}{\partial x'^{\nu}}\right) B^{\alpha}{}_{\beta}. \tag{2.31}$$

It's always good practice to check that all of the free indices match on each side. In Eq. (2.31), for example, the lefthand side contains one free index, μ, in the upper position, and one free index, ν, in the lower position; and so does the righthand side.

Tensors of higher rank transform following the pattern of Eqs. (2.29) - (2.31), namely, mutliply by one factor of $(\partial x'^{\mu}/\partial x^{\alpha})$ for every contravariant index, μ, and one factor of $(\partial x^{\beta}/\partial x'^{\nu})$ for every covariant index, ν, and then sum over all repeated indices. Thus we see the general point that not all matrices are tensors, but every tensor may be represented in a particular coordinate system in matrix form.

Consider, for example, the convenient matrix we have already been using, $\eta_{\mu\nu}$. From Eq. (2.29), we see that under a coordinate transformation, $x^{\mu} \to x'^{\mu}$, the matrix $\eta^{\mu\nu}$ should transform as

$$\eta'^{\mu\nu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\right)\left(\frac{\partial x'^{\nu}}{\partial x^{\beta}}\right) \eta^{\alpha\beta}. \tag{2.32}$$

For the particular coordinate transformation of Eq. (2.15), the 00 component of $\eta'^{\mu\nu}$ should transform as

$$\begin{aligned}
\eta'^{00} &= \left(\frac{\partial t'}{\partial x^{\alpha}}\right)\left(\frac{\partial t'}{\partial x^{\beta}}\right)\eta^{\alpha\beta} \\
&= \left(\frac{\partial t'}{\partial t}\right)\left(\frac{\partial t'}{\partial x^{\beta}}\right)\eta^{0\beta} + \left(\frac{\partial t'}{\partial x}\right)\left(\frac{\partial t'}{\partial x^{\beta}}\right)\eta^{1\beta} \\
&\qquad + \left(\frac{\partial t'}{\partial y}\right)\left(\frac{\partial t'}{\partial x^{\beta}}\right)\eta^{2\beta} + \left(\frac{\partial t'}{\partial z}\right)\left(\frac{\partial t'}{\partial x^{\beta}}\right)\eta^{3\beta} \\
&= \gamma\left[\left(\frac{\partial t'}{\partial t}\right)\eta^{00} + \left(\frac{\partial t'}{\partial x}\right)\eta^{01}\right] - \gamma v\left[\left(\frac{\partial t'}{\partial t}\right)\eta^{10} + \left(\frac{\partial t'}{\partial x}\right)\eta^{11}\right] \\
&= \gamma^2\eta^{00} + \gamma^2 v^2\eta^{11} \\
&= -\gamma^2\left(1 - v^2\right) = -1,
\end{aligned} \tag{2.33}$$

upon using the values of $\eta^{\mu\nu}$ in Eq. (2.2) and the value of γ as calculated in Eq. (2.22). Thus we have found that $\eta'^{00} = \eta^{00}$. Proceeding similarly, we find that each component in the transformed frame matches the corresponding component in the original frame, or $\eta'^{\mu\nu} = \eta^{\mu\nu}$.

We may now confirm that $\eta_{\mu\nu}$ is indeed a genuine tensor and not just an arbitrary matrix, for we know how $\eta_{\mu\nu}$ must behave. When considering the restricted coordinate transformations of special relativity of the form in Eq. (2.16), which involve only inertial relative motion, we find $(\partial x'^{\mu}/\partial x^{\nu}) = \Lambda^{\mu}{}_{\nu}$, and hence Eq. (2.32) becomes $\eta'^{\mu\nu} = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\eta^{\alpha\beta}$. But from Eq. (2.25), we immediately see that

$$\eta'^{\mu\nu} = \eta^{\mu\nu}. \tag{2.34}$$

Thus we know that $\eta^{\mu\nu}$ really does transform according to the rule of Eq. (2.29), and therefore that $\eta^{\mu\nu}$ is indeed a tensor. Not every tensor of interest, even in special relativity, will transform such that $A'^{\mu\nu} = A^{\mu\nu}$; indeed, we will see an example in the next subsection of an important tensor that has a more complicated transformation than that of $\eta^{\mu\nu}$.

Before moving on, we note the transformation property of one more type of object: a scalar, that is, an object with no free indices. Consider, for example, the dot product of two arbitrary 4-vectors, A^{μ} and B^{μ}. Following the coordinate transformation $x^{\mu} \to x'^{\mu}$, we find

$$\begin{aligned} A'_{\mu}B'^{\mu} &= \eta'_{\mu\nu}\, A'^{\mu}B'^{\nu} \\ &= \eta'_{\mu\nu}\, \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}A^{\alpha}B^{\beta} \\ &= \eta_{\mu\nu}\, \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}A^{\alpha}B^{\beta} \\ &= \eta_{\alpha\beta}\, A^{\alpha}B^{\beta} = A_{\alpha}B^{\alpha}, \end{aligned} \tag{2.35}$$

where the second line follows from the transformation rule for contravariant vectors; the third line from Eq. (2.34); and the fourth line from Eq. (2.25). Thus scalar quantities, such as dot products of vectors, are unaffected by a coordinate transformation. Put another way, the transformation of Eq. (2.16) will rotate 4-vectors within four-dimensional spacetime, and hence induce changes within given components of the 4-vectors, but will not change the vectors' overall length. Such scalar quantities, much like the specific case of $s^2 = x_{\mu}x^{\mu}$ in Eq. (2.23), are relativistic invariants.

2.3 An electromagnetic example

To get a little more practice manipulating 4-vectors and tensors before moving to general relativity, we consider a quintessential example from special relativity, involving "the electrodynamics of moving bodies." That, of course, was the title of Einstein's original paper in which he introduced special relativity, in 1905. (The paper appears in English translation, with helpful commentary, in [6].)

Maxwell's contemporary, William Thomson (later Lord Kelvin), introduced the notion of a vector potential in the 1840s — while still an undergraduate at Cambridge University. Obviously not content to stick with punting on the Cam, Thomson also introduced the differential operator "curl," or $\nabla\times$, also while an undergraduate. With those mathematical tools, he and Maxwell could represent electric and magnetic fields in terms of two types of potentials: a scalar potential, ϕ, familiar since Coulomb's day, and a 3-vector potential, $\mathbf{A}$:

$$\begin{aligned} \mathbf{E} &= -\nabla\phi - \frac{\partial}{\partial t}\mathbf{A}, \\ \mathbf{B} &= \nabla \times \mathbf{A}. \end{aligned} \tag{2.36}$$

If we continue to work with Cartesian coordinates for the three spatial dimensions, then we may represent the magnetic field vector as

$$\mathbf{B} = \hat{x}\left(\partial_y A_z - \partial_z A_y\right) + \hat{y}\left(\partial_z A_x - \partial_x A_z\right) + \hat{z}\left(\partial_x A_y - \partial_y A_x\right), \tag{2.37}$$

where $\hat{x}$ is a unit vector pointing along the x axis, and I have introduced the convenient shorthand, $\partial_y \equiv \partial/\partial y$.

We may construct a relativistic 4-vector, A^μ, from these two types of potential,

$$A^\mu = (\phi, \mathbf{A}). \tag{2.38}$$

Then the electric and magnetic fields may be combined into a single tensor, $F_{\mu\nu}$, defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2.39}$$

where as usual $A_\mu = \eta_{\mu\nu} A^\nu$ (see [1], section 2.7). From the form of Eq. (2.39), we immediately see that $F_{\mu\nu}$ vanishes when $\mu = \nu$, and that $F_{\mu\nu}$ is antisymmetric in its indices, meaning that $F_{\mu\nu} = -F_{\nu\mu}$. Other components include

$$\begin{aligned} F_{01} &= \partial_t A_x - \partial_x \phi = -E_x, \\ F_{12} &= \partial_x A_y - \partial_y A_x = B_z, \\ F_{23} &= \partial_y A_z - \partial_z A_y = B_x, \end{aligned} \tag{2.40}$$

and so on. The contravariant version, $F^{\mu\nu}$, may be found by raising each index with $\eta^{\mu\nu}$:

$$F^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}F_{\alpha\beta}, \tag{2.41}$$

and thus one finds

$$\begin{aligned} F^{01} &= \eta^{0\alpha}\eta^{1\beta}F_{\alpha\beta} = \left(-\delta^{0\alpha}\right)\left(+\delta^{1\beta}\right)F_{\alpha\beta} = (-1)(1)F_{01} = +E_x, \\ F^{12} &= \eta^{1\alpha}\eta^{2\beta}F_{\alpha\beta} = \left(+\delta^{1\alpha}\right)\left(+\delta^{2\beta}\right)F_{\alpha\beta} = +F_{12} = B_z, \end{aligned} \tag{2.42}$$

and so on. It is a worthwhile exercise to compute the remaining components! In the end one finds

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}. \tag{2.43}$$

Now consider the transformed tensor following a coordinate transformation. As always we have

$$F'^{\mu\nu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\right)\left(\frac{\partial x'^{\nu}}{\partial x^{\beta}}\right) F^{\alpha\beta}, \tag{2.44}$$

and thus

$$\begin{aligned} F'^{01} &= \left(\frac{\partial t'}{\partial x^{\alpha}}\right)\left(\frac{\partial x'}{\partial x^{\beta}}\right) F^{\alpha\beta} \\ &= \left(\frac{\partial t'}{\partial t}\right)\left(\frac{\partial x'}{\partial x^{\beta}}\right) F^{0\beta} + \left(\frac{\partial t'}{\partial x}\right)\left(\frac{\partial x'}{\partial x^{\beta}}\right) F^{1\beta} \\ &\quad + \left(\frac{\partial t'}{\partial y}\right)\left(\frac{\partial x'}{\partial x^{\beta}}\right) F^{2\beta} + \left(\frac{\partial t'}{\partial z}\right)\left(\frac{\partial x'}{\partial x^{\beta}}\right) F^{3\beta} \\ &= \gamma\left[-\gamma v F^{00} + \gamma F^{01}\right] - \gamma v\left[-\gamma v F^{10} + \gamma F^{11}\right], \end{aligned} \tag{2.45}$$

where we have assumed the coordinate transformation of Eq. (2.15). But in general $F^{00} = F^{11} = 0$ and $F^{01} = -F^{10}$, and so we find

$$F'^{01} = \gamma^2\left(1 - v^2\right) F^{01} = F^{01}, \tag{2.46}$$

or, from Eq. (2.43),

$$E'_x = E_x. \tag{2.47}$$

In a similar way, we find

$$\begin{aligned} E'_y = F'^{02} &= \left(\frac{\partial t'}{\partial x^{\alpha}}\right)\left(\frac{\partial y'}{\partial x^{\beta}}\right) F^{\alpha\beta} \\ &= \gamma\left(\frac{\partial y'}{\partial x^{\beta}}\right) F^{0\beta} - \gamma v\left(\frac{\partial y'}{\partial x^{\beta}}\right) F^{1\beta}, \end{aligned} \tag{2.48}$$

but $\partial y'/\partial x^{\beta} = \delta_{2\beta}$, and so

$$\begin{aligned} E'_y = F'^{02} &= \gamma\delta_{2\beta}F^{0\beta} - \gamma v\delta_{2\beta}F^{1\beta} \\ &= \gamma F^{02} - \gamma v F^{12} \\ &= \gamma\left(E_y - vB_z\right). \end{aligned} \tag{2.49}$$

Thus we find $F'^{02} \neq F^{02}$, or $E'_y \neq E_y$.

This is just as it should be. Consider a situation as viewed in frame S, in which within a certain region of space the electric field vanishes but the magnetic field does not: $\mathbf{E} = 0$

and $\mathbf{B} \neq 0$, with $\mathbf{B} = B_z\hat{z}$ entirely along the $\hat{z}$ direction. Then a test charge moving along the x axis with speed v, as seen in frame S, would experience a force given by $\mathbf{F} = q\mathbf{v} \times \mathbf{B} = -qvB_z\hat{y}$. Now perform a coordinate transformation to a frame moving along with the test charge. To an observer in the new frame, there would exist an electric field with nonzero component E'_y and hence the test charge would experience a force given by $\mathbf{F}' = q\mathbf{E}' = qE'_y = -q\gamma vB_z\hat{y}$, again in the $\hat{y}$ direction. Both observers would see the test charge experience a force, in the same direction. They would disagree on the magnitude of the force (their expressions differ by a factor of γ), because they would also disagree on measurements of time by the same factor of γ (see [1], section 2.3). Nonetheless, they may relate the components of their electric and magnetic field vectors by means of the tensor $F^{\mu\nu}$ and its transformation.

3 General Relativity Basics

3.1 Metric tensor and affine connection

Arguably the most important feature of special relativity is the requirement that the spacetime interval, s^2 in Eq. (2.23), remains invariant for all inertial observers. We may imagine shrinking that interval down to an infinitesimal, represented by the differential ds. Upon using Eq. (2.2), the invariant interval takes the form

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{3.1}$$

The leap that Einstein made toward general relativity was to consider arbitrary coordinate transformations rather than just the restricted class of Lorentz transformations, and to search for a comparable quantity that would remain invariant for all observers. After several years of intense effort, he made the move to generalize the tensor $\eta_{\mu\nu}$, each of whose components was simply a constant, to a new tensor $g_{\mu\nu}(x)$, whose components, in general, depend on space and time. The starting point for general relativity thus became

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu dx^\nu. \tag{3.2}$$

The tensor $g_{\mu\nu}(x)$ is known as the "metric tensor," and the expression in Eq. (3.2) is called the "line element." Remember of course that because of the Einstein summation convention, Eq. (3.2) actually stands for a whole series of terms,

$$\begin{aligned} ds^2 = g_{00}dx^0dx^0 &+ g_{01}dx^0dx^1 + g_{02}dx^0dx^2 + g_{03}dx^0dx^3 \\ &+ g_{10}dx^1dx^0 + g_{11}dx^1dx^2 + g_{12}dx^1dx^2 + g_{13}dx^1dx^3 \\ &+ g_{20}dx^2dx^0 + g_{21}dx^2dx^1 + g_{22}dx^2dx^2 + g_{23}dx^2dx^3 \\ &+ g_{30}dx^3dx^0 + g_{31}dx^3dx^1 + g_{32}dx^3dx^2 + g_{33}dx^3dx^3. \end{aligned} \tag{3.3}$$

Note that the metric tensor must be symmetrical in its indices, $g_{\mu\nu} = g_{\nu\mu}$, because there is no difference between, say, the product $dx^0 dx^1$ and $dx^1 dx^0$. Thus although the metric tensor includes 16 components (when working in four spacetime dimensions, as we are here), only 10 components are actually independent.

Just as for $\eta_{\mu\nu}$, we may also construct an inverse metric, $g^{\mu\nu}$, using the relation

$$g^{\mu\nu} g_{\mu\alpha} = \delta^{\nu}{}_{\alpha}, \tag{3.4}$$

and raise and lower indices on other objects using $g_{\mu\nu}$ and $g^{\mu\nu}$.

The fact that the metric tensor depends on space and time suggests that observers in different reference frames need to make more complicated efforts to compare their clock-rates and meter sticks than in the examples of Section 2. The difference in measured clock-rates between two observers, for example, can depend not only on their relative velocity but also on their *locations* in space. This is a first indication that general relativity concerns the bending and warping of spacetime: the means with which observers measure time intervals and spatial lengths themselves depend on the behavior of space and time in their immediate vicinity.

Because $g_{\mu\nu}$ depends on x^α, we can have (in general)

$$\frac{\partial}{\partial x^\alpha} g_{\mu\nu} \equiv \partial_\alpha g_{\mu\nu} \neq 0. \tag{3.5}$$

A particularly convenient combination of the partial derivatives of $g_{\mu\nu}$ is known as the "affine connection" or "Christoffel symbol" (see [1], section 3.3), and is given by

$$\Gamma^{\mu}{}_{\alpha\beta} \equiv \frac{1}{2} g^{\mu\lambda} \left[\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\lambda} - \partial_\lambda g_{\alpha\beta} \right]. \tag{3.6}$$

Since $g_{\mu\nu}$ is symmetric in its two indices, note that $\Gamma^{\mu}{}_{\alpha\beta}$ is symmetric in its two lower indices: $\Gamma^{\mu}{}_{\alpha\beta} = \Gamma^{\mu}{}_{\beta\alpha}$. Note further that the Christoffel symbol vanishes identically if the metric tensor is independent of space and time. That is, if each component of $g_{\mu\nu}(x)$ were simply a constant, akin to $\eta_{\mu\nu}$, then $\Gamma^{\mu}{}_{\alpha\beta} = 0$.

In order to preserve the invariance of ds^2 in Eq. (3.2), $g_{\mu\nu}(x)$ must transform as a tensor. To see this, consider the behavior of ds^2 under an arbitrary coordinate transformation, $x^\mu \to x'^\mu$:

$$ds'^2 = g'_{\mu\nu} dx'^\mu dx'^\nu = g'_{\mu\nu} \left(\frac{\partial x'^\mu}{\partial x^\alpha} \right) \left(\frac{\partial x'^\nu}{\partial x^\beta} \right) dx^\alpha dx^\beta. \tag{3.7}$$

Requiring that $ds'^2 = ds^2$ thus requires

$$g'_{\mu\nu} \left(\frac{\partial x'^\mu}{\partial x^\alpha} \right) \left(\frac{\partial x'^\nu}{\partial x^\beta} \right) = g_{\alpha\beta}, \tag{3.8}$$

or, upon rearranging the partial derivative terms,

$$g'_{\mu\nu} = \left(\frac{\partial x^\alpha}{\partial x'^\mu}\right)\left(\frac{\partial x^\beta}{\partial x'^\nu}\right) g_{\alpha\beta}, \tag{3.9}$$

which is precisely the transformation required of a covariant tensor, as in Eq. (2.30).

Whereas $g_{\mu\nu}$ is a genuine tensor, however, the Christoffel symbol defined in Eq. (3.6) is not — a helpful reminder that not all arrays or matrices actually transform as tensors. Weinberg works out the transformation explicitly in Eq. (4.5.2) of [1] (on p. 100). The point is that $\Gamma'^{\mu}{}_{\alpha\beta}$ consists of two terms, the second of which spoils the tensor transformation rule. (It is worthwhile deriving Weinberg's Eq. (4.5.2) on your own.)

Although it is not itself a tensor, $\Gamma^{\mu}{}_{\alpha\beta}$ is of great aid in constructing *other* objects that do behave as proper tensors. The first to consider is known as the "covariant derivative" (see [1], section 4.6). Physicists use many symbols to denote covariant differentiation, including

$$A^{\mu}{}_{;\nu}, \;\; \nabla_\nu A^\mu, \;\text{ and }\; \mathcal{D}_\nu A^\mu. \tag{3.10}$$

In these notes I will stick with the notation $\mathcal{D}_\nu$. For a contravariant vector, the covariant derivative is defined as

$$\mathcal{D}_\nu A^\mu \equiv \partial_\nu A^\mu + \Gamma^{\mu}{}_{\nu\beta} A^\beta, \tag{3.11}$$

whereas for a covariant vector, the covariant derivative is defined as

$$\mathcal{D}_\nu A_\mu \equiv \partial_\nu A_\mu - \Gamma^{\alpha}{}_{\nu\mu} A_\alpha. \tag{3.12}$$

It's always a good habit to check that the indices match on both sides. The pattern for covariant differentiation of a tensor follows this basic pattern, as illustrated in Weinberg's Eq. (4.6.10) in [1] (p. 104): add a factor of Γ with appropriate indices for each contravariant index of the original tensor, and subtract an appropriate factor of Γ for each covariant index of the tensor. The covariant derivative of a scalar quantity, meanwhile, is simply equal to the partial derivative:

$$\mathcal{D}_\nu F = \partial_\nu F. \tag{3.13}$$

As you should confirm for yourself by hand, the covariant derivative of a vector itself transforms as a tensor — that is the upshot of Weinberg's Eqs. (4.6.3) and (4.6.7) in [1] (pp. 103-104). Equally important, the covariant derivative of the metric tensor vanishes identically:

$$\mathcal{D}_\nu g_{\alpha\beta} = 0. \tag{3.14}$$

Indeed, the requirement that $\mathcal{D}_\nu g_{\alpha\beta} = 0$ can be used to fix the form of $\Gamma^{\mu}{}_{\alpha\beta}$. You should practice writing out all the terms implied in Eq. (3.14) to see just how the cancellations occur.

3.2 Curvature

Although covariant differentiation transforms in the way one would expect for a tensor, the application of *two* covariant derivatives to a given object does not commute. This result is not terribly surprising, since the arrays ${\Gamma^\mu}_{\alpha\beta}$ may be represented as matrices, and in general matrices do not commute. The *difference* between applying the covariant derivatives in one order minus the reverse order transforms as a tensor, and in fact can be used to define one of the most important tensors we will ever encounter in our studies of general relativity, the Riemann curvature tensor, ${R^\mu}_{\nu\lambda\sigma}$:

$$\mathcal{D}_\lambda \mathcal{D}_\sigma A^\mu - \mathcal{D}_\sigma \mathcal{D}_\lambda A^\mu = -{R^\mu}_{\nu\lambda\sigma} A^\nu, \tag{3.15}$$

or

$$\mathcal{D}_\lambda \mathcal{D}_\sigma A_\mu - \mathcal{D}_\sigma \mathcal{D}_\lambda A_\mu = {R^\nu}_{\mu\lambda\sigma} A_\nu, \tag{3.16}$$

where

$${R^\mu}_{\nu\lambda\sigma} \equiv \partial_\lambda {\Gamma^\mu}_{\nu\sigma} - \partial_\sigma {\Gamma^\mu}_{\nu\lambda} + {\Gamma^\mu}_{\rho\lambda} {\Gamma^\rho}_{\nu\sigma} - {\Gamma^\mu}_{\rho\sigma} {\Gamma^\rho}_{\nu\lambda}. \tag{3.17}$$

As Weinberg explains in section 6.2 of [1], the Riemann tensor is the *unique* tensor that may be built exclusively from $g_{\mu\nu}$ and its first two derivatives, which is also linear in the second derivatives. The Riemann tensor is defined only up to an overall sign; Weinberg chose the opposite sign to the one that is now often used in the literature (and which I have adopted here), and so again one must keep Weinberg's sign conventions in mind.

The main point is that the Riemann tensor is a particular combination of first and second partial derivatives of the metric tensor, $g_{\mu\nu}(x)$. Because of the symmetries of $g_{\mu\nu}$ and ${\Gamma^\mu}_{\alpha\beta}$ in their various indices, the Riemann tensor ${R^\mu}_{\nu\lambda\sigma}$ has many useful and important symmetries as well, which are enumerated throughout section 6.6 of [1].

Note again that if one happened to use local coordinates such that every component of the metric tensor in a particular (local) reference frame were a constant, then all the partial derivatives of $g_{\mu\nu}$ would vanish; all the Christoffel symbols ${\Gamma^\mu}_{\alpha\beta}$ would vanish; the Riemann tensor ${R^\mu}_{\nu\lambda\sigma}$ would vanish; and covariant derivatives would commute. One example of such a case is the spacetime of special relativity, in which the metric tensor $g_{\mu\nu}$ assumes the special form, $\eta_{\mu\nu}$. For more general situations, however — including classes of coordinate transformations that do not respect the limitations of the Lorentz transformations — at least some components of $g_{\mu\nu}$ will depend on x^α, and hence there will remain some residual curvature, as measured by nonzero components of ${R^\mu}_{\nu\lambda\sigma}$.

From the Riemann tensor one may construct two other important objects, by making various contractions over the original four indices. Recalling that repeated indices must be summed over, we may define the Ricci tensor as

$$R_{\nu\sigma} \equiv R^{\mu}{}_{\nu\mu\sigma}. \tag{3.18}$$

The Ricci scalar is then formed by contracting the remaining two indices,

$$R \equiv g^{\mu\nu} R_{\mu\nu}. \tag{3.19}$$

3.3 Einstein's field equations

Skipping ahead in our story somewhat — which is to say, passing lightly over the several additional years with which Einstein struggled with these geometrical objects — we come to the Einstein field equations, the governing equations of general relativity:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G\, T_{\mu\nu}. \tag{3.20}$$

The quantity on the lefthand side is often given its own special symbol, the "Einstein tensor," denoted by

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R. \tag{3.21}$$

The Einstein tensor $G_{\mu\nu}$ can be thought of as a purely geometric object; it quantifies the amount that the metric tensor, $g_{\mu\nu}(x)$, varies in space and time. It is, in other words, a measure of the curvature of spacetime. Because of the underlying symmetries of $g_{\mu\nu}$, $\Gamma^{\mu}{}_{\alpha\beta}$, and hence $R^{\mu}{}_{\nu\lambda\sigma}$, the Einstein tensor is symmetric in its two indices: $G_{\mu\nu} = G_{\nu\mu}$.

The quantity on the righthand side of Eq. (3.20) consists of a constant, $8\pi G$, multiplying an object known as the "energy-momentum tensor," $T_{\mu\nu}$. The G that appears among the constants on the righthand side is Newton's gravitational constant, the same constant that appears in Newton's famous law of gravitation, $F = GmM/r^2$. Indeed, Einstein needed to use Newton's constant so that Einstein's field equations would reduce to the Newtonian predictions in the appropriate limit.

The energy-momentum tensor, $T_{\mu\nu}$, represents how the energy and pressure of matter is distributed within spacetime. Given the symmetries of $G_{\mu\nu}$, the energy-momentum tensor, too, must be symmetric in its indices: $T_{\mu\nu} = T_{\nu\mu}$.

The upshot of Einstein's field equations can therefore be interpreted in at least two ways. Reading Eq. (3.20) from left to right, we conclude that the bending and warping of spacetime (as measured by $G_{\mu\nu}$) is determined by the distribution of matter and energy (as represented by $T_{\mu\nu}$). Or we may read in the opposite direction and conclude that the curvature of spacetime determines how matter will move.

An important point to note is that the geometrical object, $G_{\mu\nu}$, obeys what is known as a "Bianchi identity":

$$\mathcal{D}_\nu G^{\mu\nu} = \mathcal{D}_\nu \left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \right) = 0, \tag{3.22}$$

which follows identically from the particular combination of partial derivatives and Christoffel symbols out of which $R^\mu{}_{\nu\lambda\sigma}$ and its contractions are constructed. The Bianchi identity in turn imposes an important requirement upon the energy-momentum tensor. Its covariant derivative, too, must vanish:

$$\mathcal{D}_\nu T^{\mu\nu} = 0 \tag{3.23}$$

Eq. (3.23) plays the role of a conservation law, akin to (but not exactly the same as) the conservation of energy in Newtonian mechanics.

References

[1] Steven Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York: Wiley, 1972).

[2] Wolfgang Rindler, *Relativity: Special, General, and Cosmological*, 2nd ed. (New York: Oxford University Press, 2006 [2001]).

[3] Tai-Pei Cheng, *Relativity, Gravitation, and Cosmology: A Basic Introduction*, 2nd ed. (New York: Oxford University Press, 2010 [2005]).

[4] James Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 2 vols. (Oxford: Clarendon, 1873; reprint edition, New York: Dover, 1954).

[5] Bruce Hunt, *The Maxwellians* (Ithaca: Cornell University Press, 1991), appendix.

[6] John Stachel, *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*, 2nd ed. (Princeton: Princeton University Press, 2005 [1998]).