Essay on Einstein's Paper

Mateen Ulhaq

April 19, 2017

Abstract

This is an essay on Einstein's 1916 review paper, "The Foundation of the Generalized Theory of Relativity". An English translation may be read online here: [link].

Einstein begins by considering the physical motivations behind the Theory of General Relativity. He does this through descriptions of thought experiments and empirical evidence.

Next, he spends a decent amount of his paper establishing the mathematical background required by introducing tensors. He introduces the contravariant transformation law, defined in the Einstein summation convention as:

$$dx'_\sigma = \frac{\partial x'_\sigma}{\partial x_\nu} dx_\nu$$

where dx_σ and dx'_σ represent an arbitrary coordinate basis which transforms in the manner of contravariant vectors, typically upper indexed in the manner A^ν. Transformation of a general contravariant tensor over the index sets $\{i_k\}_{k=1\ldots n}$ and $\{j_k\}_{k=1\ldots n}$ is given:

$$A'^{j_1 \ldots j_n} = \frac{\partial x'_{j_1}}{\partial x_{i_1}} \, \cdots \, \frac{\partial x'_{j_n}}{\partial x_{i_n}} A^{i_1 \ldots i_n}$$

He then expresses the corresponding covariant transformation, defined in the opposite manner as:

$$A'_{j_1 \ldots j_n} = \frac{\partial x_{i_1}}{\partial x'_{j_1}} \ \cdots \ \frac{\partial x_{i_n}}{\partial x'_{j_n}} A_{i_1 \ldots i_n}$$

Mixed tensor transformations may also be performed in a similar manner, while ensuring a "conservation" of contravariant (upper) and covariant (lower) indices.

Contravariant and covariant indices hold specific meanings in different contexts. For example, they can be used to define "inverses", à la $g_{ij}g^{jk} = \delta_i^k$, or as the duals of each other.

Einstein also notes that symmetrical and anti-symmetrical tensors, which he uses later in the paper, are independent from choice of coordinates ("system of reference"). Einstein omits this proof for anti-symmetrical tensors:

$$A'^{\sigma\tau} = \frac{\partial x'_\sigma}{\partial x_\mu}\frac{\partial x'_\tau}{\partial x_\nu}A^{\mu\nu} = \frac{\partial x'_\sigma}{\partial x_\mu}\frac{\partial x'_\tau}{\partial x_\nu}(-A^{\nu\mu}) = -\frac{\partial x'_\tau}{\partial x_\mu}\frac{\partial x'_\sigma}{\partial x_\nu}A^{\mu\nu} = -A'^{\tau\sigma}$$

but it is nearly identical to the one for symmetrical tensors.

Next, Einstein walks through multiplication in tensors, which are simply combinations of indices. He also mentions the mechanic of contraction over an upper and lower index, which is simply a summation. The Einstein summation notation convention becomes useful here in defining these contractions without excessive summation signs. For instance, he forms a scalar (0 order tensor) from a (2,2) tensor through a trace-like operation,

$$A = A_{\alpha\beta}^{\alpha\beta}$$

an idea which is later used to define the scalar curvature. (In our course, we use the notation $S = \mathrm{Tr}_g \,\mathrm{Ric} = g^{ij}R_{ij} = R_j^j$.)

Then, he introduces the metric tensor $g_{\mu\nu}$ to describe the "linear element"

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

He also introduces its inverse metric tensor $g^{\mu\nu}$, which is of course defined as simply the contraction over a common index $g_{ij} g^{jk} = \delta_i^k$, analogous to the inverse matrix of linear algebra.

Then, he introduces a volume element used in his theory, $\sqrt{-g} d\tau$, which is shown to be invariant of choice of coordinates, where $g = \det g_{\mu\nu}$.

In section 9, Einstein derives the familiar geodesic equations. He does this by considering a geodesic as a path which is the stationary point on the action (i.e., extremum of the functional)

$$\int_{P_1}^{P_2} d\tau$$

where $d\tau$ is proper time, which can be related to Minkowski metric by $d\tau = dt^2 - dx^2 - dy^2 - dz^2$. (Einstein uses the notation $ds \equiv -d\tau$ so we will use this for convenience.) His motivation for doing this is the idea that objects follow geodesic paths in space-time, where proper time is locally minimized.

Then, he defines a parametrization $\lambda = \lambda(x_1, x_2, x_3, x_4)$ on any curves connecting the desired endpoints $\lambda_1 = \lambda(P_1)$ and $\lambda_2 = \lambda(P_2)$. Then, he makes the substitution $w \, d\lambda = ds$, where

$$w^2 = \left(\frac{ds}{d\lambda}\right)^2 = \left(\frac{\sqrt{g_{\mu\nu} dx_\mu dx_\nu}}{d\lambda}\right)^2 = g_{\mu\nu} \frac{dx_\mu}{d\lambda} \frac{dx_\nu}{d\lambda}$$

After all this, he takes the variation of the functional to 0:

$$0 = \delta \left\{ \int_{\lambda_1}^{\lambda_2} w \, d\lambda \right\} = \int_{\lambda_1}^{\lambda_2} \delta w \, d\lambda$$

Computing the variation on w through the product rule provides:

$$\delta w = \frac{1}{w}\left\{\frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x_\sigma}\frac{dx_\mu}{d\lambda}\frac{dx_\nu}{d\lambda}\delta x_\sigma + g_{\mu\nu}\frac{dx_\mu}{d\lambda}\frac{\delta dx_\nu}{d\lambda}\right\}$$

Here, he integrates with respect to λ, argues that the choice of δx_σ is arbitrary in the resulting functional

$$\int_{\lambda_1}^{\lambda_2} d\lambda\ \varkappa_\sigma \delta x_\sigma = 0$$

and so $\varkappa_\sigma$ must vanish to 0. This gives the geodesic equations:

$$0 = \varkappa_\sigma = \frac{d}{d\lambda}\left\{\frac{g_{\mu\nu}}{w}\frac{dx_\mu}{\partial\lambda}\right\} - \frac{1}{2w}\frac{\partial g_{\mu\nu}}{\partial x_\sigma}\frac{dx_\mu}{\partial\lambda}\frac{dx_\nu}{\partial\lambda}$$

which may be rewritten by choosing λ to be arc length of the geodesic, implying $w = \sqrt{\frac{ds}{d\lambda}} = 1$, thus giving the geodesic equations in a convenient notation:

$$g_{\alpha\sigma}\frac{d^2 x_\alpha}{ds^2} + \Gamma_{\mu\nu,\sigma}\frac{dx_\mu}{ds}\frac{dx_\nu}{ds} = 0$$

In our course, we defined a curve $c(s)$ parametrized by arc length, and took perturbations on it $c_t(s)$ in a family of curves $\mathscr{C}(s,t)$ of arc length L where $\mathscr{C}(s,0) = c(s)$ in the interval $s \in [0, L]$ and with endpoints defined in a similar manner to above, $\mathscr{C}(0,t) = P_1$ and $\mathscr{C}(L,t) = P_2$. Then, we considered the arc length functional:

$$\int_0^L \left|\frac{\partial c_t}{\partial s}\right| ds$$

and took its variation to 0 by taking the functional derivative $\frac{d}{dt}\cdot|_{t=0}$ on the perturbation parameter t. From this calculation, we too obtained the equation for geodesics.

After all this, Einstein defines a notational convenience — the Christoffel symbols of the first kind, which are given by:

$$\Gamma_{\mu\nu,\sigma} \equiv \begin{bmatrix} \mu\nu \\ \sigma \end{bmatrix} = \frac{1}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x_{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x_{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x_{\sigma}}\right)$$

He also defines Christoffel symbols of the second kind as:

$$\Gamma^{\tau}_{\mu\nu} \equiv \left\{ \begin{matrix} \mu\nu \\ \tau \end{matrix} \right\} = g^{\tau\sigma}\Gamma_{\mu\nu,\sigma}$$

In section 12, Einstein introduces the Riemann curvature tensor. He does this by defining it as a commutator of covariant derivatives:

$$\nabla_{\sigma}\nabla_{\tau}A_{\mu} - \nabla_{\tau}\nabla_{\sigma}A_{\mu} = R^{\varrho}_{\mu\sigma\tau}A_{\varrho}$$

Note that the notation Einstein uses for a covariant derivative is equivalent to the modern notation we use in this course in the following manner:

$$\nabla_{\sigma}\nabla_{\tau}A_{\mu} \equiv A_{\mu\sigma\tau}$$

By contracting the contravariant index with the 2nd covariant index of the Riemann curvature tensor, he obtains the Ricci tensor:

$$R_{\mu\nu} \equiv R^{\sigma}_{\mu\sigma\tau}$$

In Section 13, Einstein makes the realization that objects (in the form of material points) travel in paths given by geodesics, which we made reference to in our discussion on Section 9. He writes the equation of motion in familiar notation:

$$\frac{d^2x_{\tau}}{ds^2} = \Gamma^{\tau}_{\mu\nu}\frac{dx_{\mu}}{ds}\frac{dx_{\nu}}{ds}$$

He also notes that if the Christoffel coefficients vanish $\Gamma^{\tau}_{\mu\nu} = 0$, as in the case of there being no gravitational field acting upon the point mass, then the point moves uniformly in a "straight line" (in the Euclidean sense). Thus, $\Gamma^{\tau}_{\mu\nu}$ are described to be the components of the gravitational field.

In Section 14, Einstein notes that many of the equations of motion are redundant in a gravitational field free of matter ($\mathrm{Ric} = 0$), reducing them down to "10 equations of 10 quantities $g_{\mu\nu}$" through symmetry.

These can then be used to show consistency of the theory of General Relativity with Newtonian gravitation and the perihelion-motion observations of Mercury. He states that these confirm that his theory must indeed be correct.

References

[1] A. Einstein, *The Foundation of the Generalised Theory of Relativity*, 1916.

Some equations used in this essay are reproduced from Einstein's 1916 review paper, "The Foundation of the Generalised Theory of Relativity" for illustration purposes.