January 21, 2023

General Relativity, Very, Very Briefly

John D. Norton
Department of HPS
University of Pittsburgh

1. The Problem

It is very difficult to grasp the essential content of the general theory of relativity by reading Einstein's narratives on the theory and by reviewing the historical papers he wrote on the way to completion of the theory. The difficulty is that Einstein's investigations were driven by his focus on certain notions of principle: the principle of equivalence, a generalized principle of relativity, the principle of general covariance and Mach's principle. As a matter of history, these principles played a decisive role in Einstein's discovery and completion of the theory. However, their role in the logical foundations of the theory is difficulty to discern. At times, they even seem to be directly contradicted by the theory. Once Einstein had completed the theory in 1915, he did not put much effort into separating what the final theory actually says and which were the heuristics that guided him. His accounts of the theory tend to recapitulate the steps he took in devising the theory. They mix the heuristics and the final results in ways that are at best puzzling and at worst misleading. My goal here is to give a compact statement of the physical foundations of the general theory, using mathematical methods and notation that conforms with that of Einstein, but without employing the principles on which Einstein placed so much emphasis.

2. Special Relativity

The best formulation of special relativity was provided by Minkowski with his conception of a four-dimensional spacetime. The theory is captured in a few ideas:

Spacetime: Spacetime has four dimensions and events in spacetime are identified by the real valued coordinates t, x, y, z.

Line element: The interval between two neighboring events, separated by coordinate differences dt, dx, dy, dz, is

(1) $$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

If $ds^2 > 0$, then the events are timelike related and ds is (c x the proper time) elapsed between the two events, as measured by a co-moving clock.

If $ds^2 < 0$, then the events are spacelike related and $|ds|$ is the proper distance between the two events, as measured by rods.

If $ds^2 = 0$, then the events are lightlike related and can be the connected by a light signal.

Geodesics. Trajectories in spacetime that render the interval s extremal

$$\delta \int ds = 0$$

are the inertial trajectories of free bodies in spacetime, if the geodesics are timelike; and the straight lines of the Euclidean three-space, if the geodesics are spacelike.

Free bodies moving along geodesics of the line element (1) conform with

(2) $$d^2x/dt^2 = d^2y/dt^2 = d^2x/dt^2 = 0.$$

These assumptions are enough to deliver the content of special relativity. They also give us much of the content of general relativity. Only a few adjustments are needed to arrive at general relativity, as we shall see shortly. (Most importantly, the expression for the line element (1) will be generalized.)

3. Properties of Special Relativity

The familiar content of special relativity is now recoverable. For example, the time dilation effect can be read off (1). If a body is moving with three velocity $(v, 0, 0) = (dx/dt, 0, 0)$, then proper time τ dilates with respect to the time coordinate t according to

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - dx^2$$

so that $(d\tau/dt)^2 = 1 - (1/c^2)(d\mathrm{x}/dt)^2 = 1 - (v/c)^2$. We then have

$$\frac{d\tau}{dt} = \sqrt{1 - v^2/c^2}$$

If we recall that the time coordinate t coincides with the proper time of clocks that have zero coordinate velocity in this coordinate system, we see that this last result is the familiar time dilation effect.

The most important result is that the line element admits the Lorentz transformation as a symmetry. That is, if we replace the coordinates (t, x, y, z) by (t', x', y', z') where

$$t' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t - \left(\frac{v}{c^2}\right)x\right) \quad x' = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x - vt) \quad y' = y \quad z' = z$$

then the line element becomes

(1) $$ds^2 = c^2 dt'^2 - dx'^2 - dy'^2 - dz'^2$$

Crucially, the value of *ds* is unchanged. It is an invariant. This symmetry corresponds to the satisfaction of the principle of relativity of inertial motion. It is deduced and not postulated.

Trajectories of a point moving at the speed of light satisfy

$$ds = 0$$

that is

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = 0$$

If we suppress motion in the *y* and *z* direction, this is equivalent to

$$dx/dt = \pm c$$

Since the line element (1) is invariant under a Lorentz transformation, this result corresponds to the light postulate. Again, it is deduced and not postulated.

4. Introduction of Arbitrary Coordinates

Special relativity can proceed using the special coordinates (t, x, y, z) and their Lorentz transforms. However, in order to introduce curvature into the spacetime geometry, it is convenient mathematically to introduce arbitrary coordinates. That is, we introduce four new coordinates $x^i = (x^0, x^1, x^2, x^3)$ by any suitably differentiable, invertible functions of the original coordinates. The raised indices 0, 1, 2, 3 are not powers but labels. Here I differ from Einstein's notation. His labels commonly varied over 1, 2, 3, 4 and where I have a Latin index in x^i, he would use a Greek index, x^μ.) Transforming the line element into the new coordinates turns (1) into a huge expression with 16 terms:

$$ds^2 = A\,(dx^0)^2 + B\,dx^0\,dx^1 + \ldots + P\,(dx^3)^2$$

The 16 coefficients are $g_{00} = A$, $g_{01} = B$, … The expression for the line element can then be written as

(3) $$ds^2 = g_{ik}\,dx^i\,dx^k$$

Here the Einstein summation convention is used. If we repeat an index, then we sum over it. That means the expression is really 16 terms:

$$ds^2 = g_{00}\, dx^0\, dx^0 + g_{01}\, dx^0\, dx^1 + \ldots + g_{33}\, dx^3\, dx^3$$

These sixteen quantities, g_{ik}, are the components of the metric tensor and fix the geometric properties of the spacetime. We cannot have any metric tensor in special relativity. Rather we must have one that always allows us to transform back to the original simpler coordinate expression (1). Geometrically, this is the condition of flatness of the metric. A necessary and sufficient condition for flatness is the vanishing of the Riemann curvature tensor

(4) $$R^i{}_{klm} = 0$$

The Riemann curvature tensor is a complicated expression in the derivatives of the g_{ik} and the full expression can readily be found in a relativity text of your choice.

The introduction of arbitrary coordinates means that the expression for the trajectory of an inertially moving body becomes more complicated. Equation (2) above becomes

(5) $$\frac{d^2x^i}{ds^2} + \begin{Bmatrix} i \\ km \end{Bmatrix} \frac{dx^k}{ds}\frac{dx^m}{ds} = 0$$

where the Christoffel symbols of the second kind are

$$\begin{Bmatrix} i \\ km \end{Bmatrix} = \frac{1}{2} g^{ir} \left(\frac{dg_{rk}}{dx^m} + \frac{dg_{rm}}{dx^k} - \frac{dg_{km}}{dx^i} \right)$$

While equations (3), (4) and (5) look a great deal more complicated than equations (1) and (2), they are just saying the same thing, but in fancier mathematical clothing. The advantage is that now, formally, rather little new is needed to make the step to general relativity.

5. General Relativity: From Flatness to Curvature

The step to general relativity consists in just one big idea. The geometrically flat metric defined by (3) and (4) above is replaced by metrics that can be curved geometrically. The geometrical curvature corresponds to familiar gravitational effects. Free bodies will still follow the geodesics of (5). Because of the curvature of the spacetime, they will behave rather differently from special relativity. They will be, for example, the orbital motions of planets.

The important question is how the condition of flatness (4) is to be relaxed. If we consider regions of spacetime where no ordinary matter is present, then there is a very simple answer. We just need to relax the flatness condition a little. The Riemann curvature tensor $R^i{}_{klm}$ has 4 x 4 x 4 x 4 components.[1] Instead of requiring each of them individually to vanish,

[1] That is an exaggeration. Many of them are duplicates so that there are only 20 independent components.

Einstein's theory only requires that certain sums of them vanish. That is, the quantity R_{ik} now called the Ricci tensor, is set to zero:

(6) $$R_{ik} = R^m{}_{imk} = R^0{}_{i0k} + R^1{}_{i1k} + R^2{}_{i2k} + R^3{}_{i3k} = 0$$

This relaxation opens up the metric fields possible just enough to those that capture gravitational waves in empty space and gravitational fields in the empty space surrounding massive objects like our sun and the earth.

The more complicated case arises if we ask after the metric field in regions of spacetime where ordinary matter is present. In Newtonian theory, in this case, we just need to know the density of matter to determine the corresponding gravitational field. In relativity theory, no single quantity represents the gravity producing matter. Rather the metric field is affected by the mass-energy density at the event, the momentum density and energy flux at the event and by any stresses that may be present. All these quantities are captured in the stress-energy tensor, T_{ik}.

The easiest way to add it to Einstein's source free gravitational field equations (6) would merely be in a simple equality: $R_{ik} = \kappa T_{ik}$, where κ is a constant of proportionality. That simplest proposal fails since the stress-energy tensor must satisfy general relativity's analog of the conservation of energy and momentum. It is written as

$$\nabla^i T_{ik} = 0$$

The operator ∇^i is a natural generalization of the differential divergence operator, adapted to the context of curved spacetime geometries. There turns out to be an easy way to adjust the first guess of the gravitational field equations to accommodate this conservation result. It turns out that a structure now called the Einstein tensor

$$G_{ik} = R_{ik} - (1/2)g_{ik} R$$

satisfies the condition $\nabla^i G_{ik} = 0$ identically, that is, no matter what g_{ik} is. (R is the fully contracted curvature scalar.) We then arrive at Einstein's gravitational field equations of November 1915:

$$G_{ik} = \kappa T_{ik},$$

6. Summing Up

There is much more to be said about Einstein's theory. So far nothing has been said about how other matter fields, such as the electromagnetic field, might be incorporated into the theory.

However what has been covered so far is enough to pin down the essential content of Einstein's theory. Two conclusions are important:

(i) The basic structure of the theory is already present in special relativity. The major transition resides in one step: allowing the metric of spacetime to become geometrically curved.

(ii) There has been no need to introduce discussion of principles, such as the principle of equivalence, a generalized principle of relativity or Mach's principle. While some of these may be useful in understanding how Einstein came to discover general relativity and how he thought of it, they do not provide the simplest conception of the theory.