

10.3: Appendix A (Part 3)

"Does the Inertia of a Body Depend Upon its Energy Content?"

A. Einstein, Annalen der Physik. 18 (1905) 639.

The results of the previous investigation lead to a very interesting conclusion, which is here to be deduced.

I based that investigation on the Maxwell-Hertz equations for empty space, together with the Maxwellian expression for the electromagnetic energy of space, and in addition the principle that:—

The laws by which the states of physical systems alter are independent of the alternative, to which of two systems of coordinates, in uniform motion of parallel translation relative to each other, these alterations of state are referred (principle of relativity).

With these principles³² as my basis I deduced *inter alia* the following result (§ 8):—

Let a system of plane waves of light, referred to the system of coordinates (x, y, z), possess the energy l; let the direction of the ray (the wave-normal) make an angle ϕ with the axis of x of the system. If we introduce a new system of coordinates (ξ, η, ζ) moving in uniform parallel translation with respect to the system (x, y, z), and having its origin of coordinates in motion along the axis of x with the velocity v, then this quantity of light—measured in the system (ξ, η, ζ)—possesses the energy³³

$$l^* = l \frac{1 - \frac{v}{c} \cos \phi}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (10.3.1)$$

where c denotes the velocity of light. We shall make use of this result in what follows.

Notes

³² The principle of the constancy of the velocity of light is of course contained in Maxwell's equations.—AS

³³ See chapter 4 [problem 11](#).—BC

Let there be a stationary body in the system (x, y, z), and let its energy—referred to the system (x, y, z) be E_0 . Let the energy of the body relative to the system (ξ, η, ζ) moving as above with the velocity v, be H_0 .

Let this body send out, in a direction making an angle ϕ with the axis of x, plane waves of light, of energy $\frac{1}{2}L$ measured relative to (x, y, z), and simultaneously an equal quantity of light in the opposite direction. Meanwhile the body remains at rest with respect to the system (x, y, z). The principle of energy must apply to this process, and in fact (by the principle of relativity) with respect to both systems of coordinates. If we call the energy of the body after the emission of light E_1 or H_1 respectively, measured relative to the system (x, y, z) or (ξ, η, ζ) respectively, then by employing the relation given above we obtain

$$\begin{aligned} E_0 &= E_1 + \frac{1}{2}L + \frac{1}{2}L, \\ H_0 &= H_1 + \frac{1}{2}L \frac{1 - \frac{v}{c} \cos \phi}{\sqrt{1 - \frac{v^2}{c^2}}} + \frac{1}{2}L \frac{1 + \frac{v}{c} \cos \phi}{\sqrt{1 - \frac{v^2}{c^2}}} \\ &= H_1 + \frac{L}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned}$$

By subtraction we obtain from these equations

$$H_0 - E_0 - (H_1 - E_1) = L \left\{ \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right\}. \quad (10.3.2)$$

The two differences of the form $H - E$ occurring in this expression have simple physical significations. H and E are energy values of the same body referred to two systems of coordinates which are in motion relative to each other, the body being at rest in one of the two systems (system (x, y, z)). Thus it is clear that the difference $H - E$ can differ from the kinetic energy K of the body, with

respect to the other system (ξ, η, ζ) , only by an additive constant C , which depends on the choice of the arbitrary additive constants³⁴ of the energies H and E . Thus we may place

$$\begin{aligned} H_0 - E_0 &= K_0 + C, \\ H_1 - E_1 &= K_1 + C, \end{aligned}$$

since C does not change during the emission of light. So we have

$$K_0 - K_1 = L \left\{ \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right\}. \quad (10.3.3)$$

Note

A potential energy U is only defined up to an additive constant. If, for example, U depends on the distance between particles, and the distance undergoes a Lorentz contraction, there is no reason to imagine that the constant will stay the same.—BC

The kinetic energy of the body with respect to (ξ, η, ζ) diminishes as a result of the emission of light, and the amount of diminution is independent of the properties of the body. Moreover, the difference $K_0 - K_1$, like the kinetic energy of the electron (§ 10), depends on the velocity.

Neglecting magnitudes of fourth and higher orders³⁵ we may place

$$K_0 - K_1 = \frac{1}{2} \frac{L}{c^2} v^2. \quad (10.3.4)$$

Note

The purpose of making the approximation is to show that under realistic lab conditions, the effect exactly mimics a change in Newtonian mass.

From this equation it directly follows³⁶ that:—

If a body gives off the energy L in the form of radiation, its mass diminishes by $\frac{L}{c^2}$. The fact that the energy withdrawn from the body becomes energy of radiation evidently makes no difference, so that we are led to the more general conclusion that

The mass of a body is a measure of its energy-content; if the energy changes by L , the mass changes in the same sense by $\frac{L}{9} \times 10^{20}$, the energy being measured in ergs, and the mass in grammes.

It is not impossible that with bodies whose energy-content is variable to a high degree (e.g. with radium salts) the theory may be successfully put to the test.

If the theory corresponds to the facts, radiation conveys inertia between the emitting and absorbing bodies.

Note

The object has the same velocity v before and after emission of the light, so this reduction in kinetic energy has to be attributed to a change in mass.—BC

"The Foundation of the General Theory of Relativity"

A. Einstein, Annalen der Physik 49 (1916) 769.

[A one-page introduction relating to history and personalities is omitted.—BC]

A. Fundamental Considerations on the Postulate of Relativity

§1. Observations on the Special Theory of Relativity

The special theory of relativity is based on the following postulate, which is also satisfied by the mechanics of Galileo and Newton. If a system of coordinates K is chosen so that, in relation to it, physical laws hold good in their simplest form, the *same* laws also hold good in relation to any other system of coordinates K' moving in uniform translation relative to K . This postulate we call the "special principle of relativity." The word "special" is meant to intimate that the principle is restricted to the case when K' has a

motion of uniform translation³⁷ relative to K , but that the equivalence of K' and K does not extend to the case of non-uniform motion of K' relative to K .

Note

Here Einstein defines the distinction between special and general relativity according to whether accelerated frames of reference are allowed. The modern tendency is to pose this distinction in terms of flat versus curved spacetime, so that accelerated frames of reference in flat spacetime are considered to be part of special relativity. None of this has anything to do with the ability to describe accelerated *objects*. For example, special relativity is perfectly capable of describing the twin paradox.—BC

Thus the special theory of relativity does not depart from classical mechanics through the postulate of relativity, but through the postulate of the constancy of the velocity of light *in vacuo*, from which, in combination with the special principle of relativity, there follow, in the well-known way, the relativity of simultaneity, the Lorentzian transformation and the related laws for the behaviour of moving bodies and clocks.

The modification to which the special theory of relativity has subjected the theory of space and time is indeed far-reaching, but one important point has remained unaffected. For the laws of geometry, even according to the special theory of relativity, are to be interpreted directly as laws relating to the possible relative positions of solid bodies at rest; and, in a more general way, the laws of kinematics are to be interpreted as laws which describe the relations of measuring bodies and clocks. To two selected material points of a stationary rigid body there always corresponds a distance of quite definite length, which is independent of the locality and orientation of the body, and is also independent of the time. To two selected positions of the hands of a clock at rest relative to the privileged system of reference there always corresponds an interval of time of a definite length, which is independent of place and time. We shall soon see that the general theory of relativity cannot adhere to this simple physical interpretation of space and time.³⁸

Note

Einstein is just starting to lay out his argument, and has not yet made clear in what sense these statements about location-independence of clocks and rulers could be empirically tested. It becomes more clear later that he means something like this. We could try to fill spacetime with a lattice of clocks and rulers, to synchronize the clocks, and to construct the lattice so that it consisted of right angles and equal-length line segments. This succeeds in special relativity, so that the geometry of spacetime is compatible with frames of reference that split up spacetime into 3+1 dimensions, where the three dimensions are Euclidean. The same prescription fails in general relativity.—BC

§2. The Need for an Extension of the Postulate of Relativity

In classical mechanics, and no less in the special theory of relativity, there is an inherent epistemological defect which was, perhaps for the first time, clearly pointed out by Ernst Mach. We will elucidate it by the following example:³⁹ — Two fluid bodies of the same size and nature hover freely in space at so great a distance from each other and from all other masses that only those gravitational forces need be taken into account which arise from the interaction of different parts of the same body. Let the distance between the two bodies be invariable, and in neither of the bodies let there be any relative movements of the parts with respect to one another.

Note

This example was described in [section 3.6](#).—BC

But let either mass, as judged by an observer at rest relative to the other mass, rotate with constant angular velocity about the line joining the masses. This is a verifiable relative motion of the two bodies. Now let us imagine that each of the bodies has been surveyed by means of measuring instruments at rest relative to itself, and let the surface of S_1 prove to be a sphere, and that of S_2 an ellipsoid of revolution. Thereupon we put the question — What is the reason for this difference in the two bodies? No answer can be admitted as epistemologically satisfactory,⁴⁰ unless the reason given is an *observable fact of experience*. The law of causality has not the significance of a statement as to the world of experience, except when *observable facts* ultimately appear as causes and effects.

Note

Of course an answer may be satisfactory from the point of view of epistemology, and yet be unsound physically, if it is in conflict with other experiences. —AS

Newtonian mechanics does not give a satisfactory answer to this question. It pronounces as follows: — The laws of mechanics apply to the space R_1 , in respect to which the body S_1 is at rest, but not to the space R_2 , in respect to which the body S_2 is at rest. But the privileged space R_1 of Galileo, thus introduced, is a merely *factitious*⁴¹ cause, and not a thing that can be observed. It is therefore clear that Newton's mechanics does not really satisfy the requirement of causality in the case under consideration but only apparently does so, since it makes the factitious cause R_1 responsible for the observable difference in the bodies S_1 and S_2 .

Note

i.e., artificial —BC

The only satisfactory answer must be that the physical system consisting of S_1 and S_2 reveals within itself no imaginable cause to which the differing behaviour of S_1 and S_2 can be referred. The cause must therefore lie *outside* this system. We have to take it that the general laws of motion, which in particular determine the shapes of S_1 and S_2 , must be such that the mechanical behaviour of S_1 and S_2 is partly conditioned in quite essential respects, by distant masses which we have not included in the system under consideration. These distant masses and their motions relative to S_1 and S_2 must then be regarded as the seat of the causes (which must be susceptible to observation) of the different behaviour of our two bodies S_1 and S_2 . They take over the role of the factitious cause R_1 . Of all imaginable spaces R_1 , R_2 , etc., in any kind of motion relative to one another there is none which we may look upon as privileged a priori without reviving the above-mentioned epistemological objection. *The laws of physics must be of such a nature that they apply to systems reference in any kind of motion.*⁴² Along this road we arrive at an extension at the postulate of relativity.

Note

At this time, Einstein had high hopes that his theory would be fully Machian. He was already aware of the Schwarzschild solution (he refers to it near the end of the paper), which offended his Machian sensibilities because it imputed properties to spacetime in a universe containing only a single point-mass. In the present example of the bodies S_1 and S_2 , general relativity actually turns out to give the non-Machian result which Einstein here says would be unsatisfactory.—BC

In addition to this weighty argument from the theory of knowledge, there is a well-known physical fact which favours an extension of the theory of relativity. Let K be a Galilean system of reference, i.e., a system relative to which (at least in the four-dimensional region under consideration) a mass, sufficiently distant from other masses, is moving with uniform motion in a straight line. Let K' be a second system of reference which is moving relative to K in *uniformly accelerated* translation. Then, relative to K' , a mass sufficiently distant from other masses would have an accelerated motion such that its acceleration and direction of acceleration are independent of the material composition and physical state of the mass.

Does this permit an observer at rest relative to K' to infer that he is on a “really” accelerated system of reference? The answer is in the negative; for the above-mentioned relation of freely movable masses to K' may be interpreted equally well in the following way. The system of reference K' is unaccelerated, but the space-time territory in question is under the sway of a gravitational field, which generates the accelerated motion of the bodies relative to K' .

This view is made possible for us by the teaching of experience as to the existence of a field of force, namely, the gravitational field, which possesses the remarkable property of imparting the same acceleration to all bodies.⁴³ The mechanical behaviour of bodies relative to K' is the same as presents itself to experience in the case of systems which we are wont to regard as “stationary” or as “privileged.” Therefore, from the physical standpoint, the assumption readily suggests itself that the systems K and K' may both with equal right be looked upon as “stationary” that is to say, they have an equal title as systems of reference for the physical description of phenomena.

Note

Eötvös has proved experimentally that the gravitational field has this property in great accuracy.—AS

It will be seen from these reflections that in pursuing the general theory of relativity we shall be led to a theory of gravitation, since we are able to “produce” a gravitational field merely by changing the system of coordinates. It will also be obvious that the

principle of the constancy of the velocity of light *in vacuo* must be modified, since we easily recognize that the path of a ray of light with respect to K' must in general be curvilinear, if with respect to K light is propagated in a straight line with a definite constant velocity.

§3. The Space-Time Continuum. Requirement of General Covariance for the Equations Expressing General Laws of Nature

In classical mechanics, as well as in the special theory of relativity, the coordinates of space and time have a direct physical meaning. To say that a point-event has the X_1 coordinate x_1 means that the projection of the point-event on the axis of X_1 , determined by rigid rods and in accordance with the rules of Euclidean geometry, is obtained by measuring off a given rod (the unit of length) x_1 times from the origin of coordinates along the axis of X_1 . To say that a point-event has the X_4 coordinate $x_4 = t$, means that a standard clock, made to measure time in a definite unit period, and which is stationary relative to the system of coordinates and practically coincident in space with the point-event,⁴⁴ will have measured off $x_4 = t$ periods at the occurrence of the event.

Note

We assume the possibility of verifying “simultaneity” for events immediately proximate in space, or — to speak more precisely — for immediate proximity or coincidence in space-time, without giving a definition of this fundamental concept.—AS

This view of space and time has always been in the minds of physicists, even if, as a rule, they have been unconscious of it. This is clear from the part which these concepts play in physical measurements; it must also have underlain the reader’s reflections on the preceding paragraph for him to connect any meaning with what he there read. But we shall now show that we must put it aside and replace it by a more general view, in order to be able to carry through the postulate of general relativity, if the special theory of relativity applies to the special case of the absence of a gravitational field.

In a space which is free of gravitational fields we introduce a Galilean system of reference $K(x, y, z, t)$, and also a system of coordinates $K'(x', y', z', t')$ in uniform rotation⁴⁵ relative to K . Let the origins of both systems, as well as their axes of Z , permanently coincide. We shall show that for a space-time measurement in the system K' the above definition of the physical meaning of lengths and times cannot be maintained. For reasons of symmetry it is clear that a circle around the origin in the X, Y plane of K may at the same time be regarded as a circle in the X', Y' plane of K' . We suppose that the circumference and diameter of this circle have been measured with a unit measure infinitely small compared with the radius, and that we have the quotient of the two results. If this experiment were performed with a measuring-rod⁴⁶ at rest relative to the Galilean system K , the quotient would be π . With a measuring-rod at rest relative to K' , the quotient would be greater than π . This is readily understood if we envisage the whole process of measuring from the “stationary” system K , and take into consideration that the measuring-rod applied to the periphery undergoes a Lorentzian contraction, while the one applied along the radius does not.⁴⁷ Hence Euclidean geometry does not apply to K' . The notion of coordinates defined above, which presupposes the validity of Euclidean geometry, therefore breaks down in relation to the system K' . So, too, we are unable to introduce a time corresponding to physical requirements in K' , indicated by clocks at rest relative to K' . To convince ourselves of this impossibility, let us imagine two clocks of identical constitution placed, one at the origin of coordinates, and the other at the circumference of the circle, and both envisaged from the “stationary” system K . By a familiar result of the special theory of relativity, the clock at the circumference — judged from K — goes more slowly than the other, because the former is in motion and the latter at rest. An observer at the common origin of coordinates, capable of observing the clock at the circumference by means of light, would therefore see it lagging behind the clock beside him. As he will not make up his mind to let the velocity of light along the path in question depend explicitly on the time, he will interpret his observations as showing that the clock at the circumference “really” goes more slowly than the clock at the origin. So he will be obliged to define time in such a way that the rate of a clock depends upon where the clock may be.

Notes

⁴⁵ This example of a rotating frame of reference was discussed in [section 3.5](#).—BC

⁴⁶ Einstein implicitly assumes that the measuring rods are perfectly rigid, but it is not obvious that this is possible. This issue is discussed in [section 3.5](#).—BC

⁴⁷ As described in [section 3.5](#), Ehrenfest originally imagined that the circumference of the disk would be reduced by its rotation. His argument was incorrect, because it assumed the ability to start the disk rotating when it had originally been at rest. The present paper marks the first time that Einstein asserted the opposite, that the circumference is increased.—BC

We therefore reach this result: — In the general theory of relativity, space and time cannot be defined in such a way that differences of the spatial coordinates can be directly measured by the unit measuring-rod, or differences in the time coordinate by a standard clock.

The method hitherto employed for laying coordinates into the space-time continuum in a definite manner thus breaks down, and there seems to be no other way which would allow us to adapt systems of coordinates to the four-dimensional universe so that we might expect from their application a particularly simple formulation of the laws of nature. So there is nothing for it but to regard all imaginable systems of coordinates, on principle, as equally suitable for the description of nature.⁴⁸ This comes to requiring that:

*The general laws of nature are to be expressed by equations which hold good for all the systems of coordinates, that is, are covariant with respect to any substitutions whatever (generally covariant).*⁴⁹

Notes

⁴⁸ This is a conceptual leap, not a direct inference from the argument about the rotating frame. Einstein started thinking about this argument in 1912, and concluded from it that he should base a theory of gravity on non-Euclidean geometry. Influenced by Levi-Civita, he tried to carry out this project in a coordinate-independent way, but he failed at first, and for a while explored a theory that was not coordinate-independent. Only later did he return to coordinate-independence. It should be clear, then, that the link between the rotating-frame argument and coordinate-independence was not as clearcut as Einstein makes out here, since he himself lost faith in it for a while.—BC

⁴⁹ In this book I've used the more transparent terminology "coordinate independence" rather than "general covariance."—BC

It is clear that a physical theory which satisfies this postulate will also be suitable for the general postulate of relativity.⁵⁰ For the sum of all substitutions in any case includes those which correspond to all relative motions of three-dimensional systems of coordinates. That this requirement of general covariance, which takes away from space and time the last remnant of physical objectivity,⁵¹ is a natural one, will be seen from the following reflection. All our space-time verifications invariably amount to a determination of space-time coincidences.⁵² If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measurements are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hands of a clock and points on the clock-dial, and observed point-events happening at the same place at the same time.

Notes

⁵⁰ For more on this point, see [section 3.7](#).—BC

⁵¹ This is an extreme interpretation of general covariance, and one that Einstein himself didn't hew closely to later on. He presented an almost diametrically opposed interpretation in a philosophical paper, "On the aether," Schweizerische naturforschende Gesellschaft 105 (1924) 85.—BC

⁵² i.e., what this book refers to as incidence measurements ([section 3.4](#))—BC

The introduction of a system of reference serves no other purpose than to facilitate the description of the totality of such coincidences. We allot to the universe four space-time variables x_1, x_2, x_3, x_4 in such a way that for every point-event there is a corresponding system of values of the variables $x_1 \dots x_4$. To two coincident point-events there corresponds one system of values of the variables $x_1 \dots x_4$, i.e., coincidence is characterized by the identity of the coordinates. If, in place of the variables $x_1 \dots x_4$, we introduce functions of them, x'_1, x'_2, x'_3, x'_4 , as a new system of coordinates, so that the systems of values are made to correspond to one another without ambiguity, the equality of all four coordinates in the new system will also serve as an expression for the space-time coincidence of the two point-events. As all our physical experience can be ultimately reduced to such coincidences, there is no immediate reason for preferring certain systems of coordinates to others, that is to say, we arrive at the requirement of general covariance.

Contributors and Attributions

This page titled [10.3: Appendix A \(Part 3\)](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Benjamin Crowell](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.