

FROM THE "USPEKHI FIZICHESKIKH NAUK" ARCHIVE

PACS number: **04.20.–q**

Einstein's general principle of relativity

V K Frederiks

Einstein's paper under the title "Die formale Grundlagen der allgeminer Relativitätstheorie" (Formal principles of the general theory of relativity) [1] appeared in the records of proceedings of Berlin Academy of Sciences in 1914 and should be considered as his first fundamental work on the principle of relativity. Later in 1916, this paper, somewhat amended and completed, was published in *Annalen der Physik*. The reprints of this work were offered for sale. Owing to this fact, the work of Einstein became especially popular [2]. H Lorentz, who in 1915 – 1916 delivered lectures on relativity in Leiden, named them *On Einstein Theory of Gravitation* [3]; mathematician D Hilbert called his papers, appearing in 1915 – 1916, "Die Grundlagen der Physik" (Basics of physics) [4]; at last, mathematician H Weyl issued a book in 1918, dedicated to these theories, under the title *Raum, Zeit, Materie* (Space, Time, Matter) [5]. Already these titles show clear enough that the theory created by Einstein embraces the whole physics, and theories of such a kind could not but arise deep, exciting interest; this is confirmed by the fact that from the very moment of its appearance Einstein's theory became the subject of consideration for such outstanding physicists and mathematicians as Lorentz, Hilbert, and Weyl. For a more or less complete exposition, this theory requires rather subtle mathematical tools, hardly understandable for almost all the physicists. As to the popular expositions, however brilliantly written, they are good for nothing except vague, loose and dissolving views for those who would like to receive something more than a general notion on Einstein's theory.

The proposed paper is too concise to lay claim to any exhaustive explanation of Einstein's theory. Its aim is to elucidate the main concepts and their application to the solution of two or three comparatively simple questions such as, for example, questions on the motion of Mercury's perihelion or the deviation of a light ray in the gravitational field of the Sun, about which recently there was much ado. It goes without saying that the main Einsteinian concepts should not be considered as theorems which could be derived in a purely deductive way from other statements which are beyond any doubt. The elucidation of the fundamentals of the theory comes to the explanation, or better to say, to the enumeration of the reasons why the fundamentals should be considered as such. The proof of the validity of Einstein's theory should be looked for not *a priori*, but *a posteriori*. Not experimental validation of the conclusions and the foreseeing of new, unknown phenomena are the most important thing in Einstein's theory. The fundamentals of Einstein's theory are of great importance ***in principle***, and here one should seek the main value of the theory but not in several experiments strengthening Einstein's theory, however brilliant these experiments may be.

Geometry and physics. Prior to Einstein, the geometry and physics were considered as two essentially quite different sciences. In physics they considered geometry as something external in respect to physics; really the content of physics was only given by experiment. The Euclidean geometry of 3-dimensional space was only a frame, true enough necessary — because any physical phenomenon took place in this space — but in any case having nothing in common with the phenomenon. True, in the case of the now so-called 'special' principle of relativity (1905) H Minkowski used a geometry of 4-dimensional space without any signs of Euclidean geometry and connected with physics via one constant it included, which is equal to the speed of light. In Minkowski geometry the element of length was defined as $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\,\mathrm{d}t^2$, where x, y, z are properly space coordinates, t is time, and c is the speed of light. This geometry is non-Euclidean, because in Euclidean geometry one would have $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \mathrm{d}c^2 t^2$; moreover, it was connected with physics, because it included the constant c — the speed of light; nevertheless Minkowski geometry was considered as purely formal just in the same way as in physics one is viewing $\sqrt{-1}$, but more an intimate connection between physics and geometry did not exist yet.

Geometry was for physics a kind of a frame, something external, outside in respect to the content of physics. On the contrary, for some geometers physics seemed sometimes to be a science whose experimental data were necessary for consolidation of the very foundations of geometry. It is self-evident that to come to know the particulars of the foundations of geometry is beyond the scope of the present paper. In his book *La Science et l'Hypothese* (The Science and Hypothesis) [6] Poincaré gives an excellent universally understandable analysis of what refers to the foundations of geometry, how one has to consider the axioms of Euclidean geometry as well as the geometries of Lobachevsky, Riemann and all other numberless non-Euclidean geometries. We shall restrict ourselves here only to the discussion of the association between an experiment and axioms or geometrical theorems

[1] *Sitzungsberichte der Preussischen Akademie der Wissenschaften* T. XLI (Berlin, 1914) [Translated into Russian in *Sobranie Nauchnykh Trudov* (Collected Works) Vol. 1 (Eds I E Tamm, Ya A Smorodinskiĭ, B G Kuznetsov) (Moscow: Nauka, 1965) p. 326].

[2] In 1920 it was also reproduced in a small book published by Teubner-Verlag under the title: H A Lorentz, A Einstein, H Minkowski: *Das Relativitätsprinzip*.

[3] Lorentz H A *On Einstein Theory of Gravitation* (Amsterdam, 1916).

[4] *Nachrichten der Königl. Gesellchaft der Wissenschaften* (Göttingen,1915, 1916).

[5] Weyl H *Raum, Zeit, Materie* 5. gearb. Ausgabe, 1923.

[6] Poincaré H *La science et l'hypothese*.

V K Frederiks
Article first published in *Uspekhi Fizicheskikh Nauk* **2** (2) 162 (1921).

Uspekhi Fizicheskikh Nauk **169** (12) 1339 – 1350 (1999)
Translated by Yu A Danilov; edited by A Radzig

derived on the basis of these axioms. The elucidation of this point is of importance for comprehension of the Einsteinian viewpoint on geometry.

As long as there was only one geometry — that of Euclid, there was no doubt in the 'physical' truthfulness of its axioms, though as far back as Gauss considered it necessary to test directly the proposition according to which the sum of angles in a triangle is equal to two right angles. Since the advent of the geometries of Lobachevsky, Riemann and others the problem of an experimental test of geometry has become of especial importance. As is known, Lobachevsky geometry denies the Euclidean postulate, according to which through a given point one can draw only one straight line parallel to a given straight line, and offers a counterpostulate: one can draw as much such parallels as one wishes. The so-called spherical Riemann geometry deviates in the opposite direction from Euclidean geometry and completely denies the possibility of the existence of parallel straight lines. Both Lobachevsky and Riemann (in his spherical geometry) accept all other Euclidean axioms. As it is very nicely and simply said in Poincaré's book mentioned above, both geometries are logically quite possible and do not contain any intrinsic contradictions. Neither Lobachevsky, nor Riemann find the sum of the angles of triangle to be equal to two right angles. In Lobachevsky's geometry the sum of the angles was less than two right angles, in Riemann's geometry the sum was more than two right angles. Gauss in his experimental test found that to within observational error the sum of the angles of a triangle turned out to be equal to two right angles. Because the angles can be measured with a high precision, it might appear at first sight that Gauss's experiment showed that real 'physical' space (by 'physical' space one understands the space where all physical phenomena take place, in contrast to spaces which we can imagine or construct logically) is the so habitual for us normal Euclidean space. But, first, the deviations from Euclidean geometry may be so small that, despite their existence and Gauss's observations being comparatively precise, they nevertheless cannot be detected specifically in this experiment. Second, even if the experiment would give with absolute accuracy that the sum of the angles of triangle is equal to two right angles — even in this case it would be impossible to assert that physical space is Euclidean without mentioning one circumstance which is important in principle. Really, let us suppose that the experiment also brings to the sum less than two right angles. Could a physicist draw from this fact a conclusion that Euclidean geometry is not valid? First of all he would ask how the measuring of angles was carried out. The answer would be — by counting the divisions on the circular limb and by using the telescope. The latter means that the light beam is used as a straight line between two vertices of a triangle whose angles are to be measured and summed up, and the deviation from two right angles for the sum physicist could interpret, if he liked, not as the invalidity of Euclidean geometry but just as 'bending' of the light beam (on the contrary, in the case of the sum equal to two right angles, a scientist at all price upholding the viewpoint of Lobachevsky geometry could also explain the deviation from the corresponding Lobachevsky theorem as 'bending' of the light beam). But a physicist speaking of bending of a ray means that in some or other way the bending under discussion can be detected; to this end he has to use some other 'physical' apparatus which would produce a 'real', ***in his opinion***, straight line; comparing the light beam with this straight line, he could show that the ray is really bent and that his new measuring apparatus gives the sum equal to two right angles. But his triumph would be very superficial and short-term; the scientist upholding the Lobachevsky viewpoint would ask him ***to prove*** that his new apparatus produces a straight line, and without inventing some other new apparatus our physicist in no way could make such thing. And because it is impossible to invent new apparatus without end, it is clear that an experiment can answer our question only as far as we ascribe to our main apparatus, say, the light beam, the properties of a straight line. But ascribing the properties of straight line precisely to the light beam and not to some other thing depends exclusively on our will. We emphasize this fact because the geometry which Einstein makes use of — is non-Euclidean, and it may seem that the validity or nonvalidity of Einstein's theory serves as a proof of the validity or non-validity of Euclidean geometry. Meanwhile this is not so; those who want to consider Euclidean geometry to be something exclusive can continue to do so without being disturbed by Einstein's arguments and theories, *but in this case he should refuse to consider as straight lines those which are produced by our main measuring instruments*: *the light beam, the edge of a ruler, and so on*. As it will be shown, if one takes a light beam or the edge of a ruler as a straight line, then observations more precise than those carried out by Gauss will reveal deviations from Euclidean geometry.

But independently of the results given or possibly given by observations that are really carried out, it is important ***in principle*** to establish that if a straight line is physically defined, say, using the light beam, ***then only experiment can indicate us what kind of geometry is valid for the physical space***. But there are infinitely many geometries; how one can experimentally test them, which conclusions and propositions of geometries could be tested better and more conveniently? There are only a few geometries which allow the displacement of unchangeable figures (the transfer of a figure from one place in space to another; the displacement or movement of figures; the existence of a rigid body); the main of them are the geometries of Lobachevsky, Riemann and Euclid. The main propositions of these geometries were discussed by Hermann Helmholtz, Sophus Lie, Bertrand Russell and others. The so-called Riemann geometries include a much wider class of geometries. As a starting point for each of his geometries Riemann makes use of a definition of the element of length.

Let us have an n-dimensional space, $x_1, x_2, \ldots, x_n$ be n coordinates defining the position of a point in this space, and let $\mathrm{d}s$ be an element of arc length. The expression

$$\mathrm{d}s^2 = \sum_{ik} a_{ik}\,\mathrm{d}x_i\,\mathrm{d}x_k\,,$$

where a_{ik} are functions of $x_1, x_2, \ldots, x_n$, is characteristic for the geometry under consideration. For each given geometry the functions a_{ik} are of some quite definite form. For example, in the case of Euclidean geometry in 3-dimensional space

$$\mathrm{d}s^2 = \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2\,;$$

for the Lobachevsky geometry

$$\mathrm{d}s^2 = \frac{\mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2}{1 - (a^2/4)(x^2 + y^2 + z^2)}\,,$$

where a is some constant.

Putting in general $ds^2 = \sum_{ik} a_{ik}\, dx_k$, we reject not only an Euclidean axiom on parallels but some other axioms as well; thus, in the most general case the axioms which allow the possibility of transferring and superposition of figures cease to be valid (these axioms are not stated explicitly by Euclid; the details on these items can be found in "Die Grundlagen der Geometrie" by Hilbert [7]).

According to Einstein, this relation, fundamental for every geometry, should be tested experimentally. Of course, only an ideal imaginary experiment is possible here; really the test is carried out not over the expression for ds^2 itself, but over the conclusions which can be drawn from it. If the test gives for 3-dimensional space $a_{ik} = 0$ for $i \neq k$ and $a_{ik} = 1$ for $i = k$, then we have Euclidean geometry; if a_{ik} turn out to be functions of x_i, then we have a geometry whose character depends on the form of these functions.

In an Einsteinian 'special' principle of relativity one should consider time as a quantity closely linked with space measurements, inseparable from them. Naturally, Einstein takes as the basis for his new, more general theories not 3-dimensional space, but 4-dimensional space where one coordinate is time. Each physical phenomenon is determined by the position where it takes place (three space coordinates) and the moment when it happens (time coordinate); the expression for the element of length consists of increments of these four coordinates:

$$ds^2 = \sum_{ik} a_{ik}\, dx_i\, dx_k\,, \qquad i,k = 1,2,3,4\,;$$

here all four coordinates play the same role; really time (say, x_4) is not identical to space coordinates x_1, x_2, x_3. Hilbert [8] derived the conditions for a_{ik} to be satisfied in order that the fourth coordinate — time — should not lose its distinctive special signs which it should reserve in any theory.

The first Einstein fundamental proposition. Thus, the first proposition of Einstein's theory runs as follows:

T h e e l e m e n t o f l e n g t h i s d e t e r m i n e d a s

$$ds^2 = \sum_{ik} a_{ik}\, dx_i\, dx_k\,, \qquad i,k = 1,2,3,4\,,$$

a n d t h e e x p e r i m e n t s e t t l e s w h a t t h e f u n c t i o n s a_{ik} a r e.

Geometry and mechanics. Newtonian mechanics ascribes quite exclusive significance to rectilinear and uniform motion. If instead of a given coordinate system we take another one moving in respect to the first system rectilinearly and uniformly, then this change cannot be found by a direct physical experiment. Any other motion can be revealed by an experiment, because new forces would appear which did not exist before. Let us consider a very simple case: motion along a circle. This case gives us an opportunity to indicate two remarkable facts: one of which is very important in principle and the other — from the experimental viewpoint.

Let us take a physical body having the form of a sphere, and suppose that in the whole of space there are no other bodies. Can we solve the question of whether this sphere rotates or not? ***Beyond*** the sphere there is no physical body, no physical point which could help us in this respect. We have to consider what happens on the very surface of the sphere or inside it. Let us suppose that we found on the surface the presence of some centrifugal force, took notice that the sphere is somewhat flattened at the poles and that the plane of oscillation of a Foucault pendulum rotates. This will compel us to suppose that our sphere is rotating, we even shall calculate the velocity of its rotation. But then will come the question ***with respect to what does it rotate***, because there are no external bodies in respect to which it could rotate. Evidently, there must be some space which itself is not distinguished by anything, does not contain any physical body and for this reason is not accessible by itself for observation; it is in this 'absolute' nonphysical space the rotation of our sphere takes place. But everything which does not have physical reality and for this reason is inaccessible to physical observation refers, if you like, to metaphysics but in no way to physics. One can believe or not believe in such an absolute space, but it is impossible to use it as an object which exists really, physically. But then one has to say ***that Newtonian mechanics can answer a question which in essence cannot be answered 'physically'***. It is a paradox, to which attention for the first time was drawn by E Mach; Einstein drove this paradox out of oblivion and gave the answer: Newtonian mechanics, in general, is not valid. Real mechanics gives for such a sphere neither a centrifugal force, nor a movement of the plane of oscillation for a Foucault pendulum, etc. — all these forces and phenomena appear only when the rotation of our body takes place in respect to some other 'physical' space which can be detected with the physical bodies which it contains. Rectilinear and uniform motion as well as circular or some other motion does not play any exclusive role; all coordinate systems and all possible their displacements are of equal worth. If there is only one sphere and nothing else, then we can assert that it rotates or rests, jumps or moves as it likes. No physical phenomena can reveal this, because all these jumps, rotations, etc. take place not in respect to 'physical' space, but ***can only be thought*** in respect to other space which does not really exist. To show the possibility of such mechanics was among the greatest services of A Einstein.

The other remarkable fact stemming from observations of rotation is that the centrifugal force is always proportional to the mass of the rotating body. According to the second Newton's law, gravitational force is also proportional to mass, but in the formula of Newton's law mass plays the role of the cause creating gravity; meanwhile, in the formula of centrifugal force engendered by rotation, it plays quite a passive role; mass actively creating a force and inert or passive mass — just a numerical coefficient — under experimental tests turn out to coincide with high accuracy [9]. This fact cannot be accidental, but Newtonian mechanics does not explain it. Newton in his second law of motion just put forward, like a postulate, the requirement according to which the inert mass times acceleration should be equal to the force — in particular, to the force produced by the same mass, the same quantitatively, but already operating actively.

The identity of masses — active or gravitating and passive or inert — Einstein raises to a principle and calls it the ***principle of equivalence***.

[7] Hilbert D "Die Grundlagen der Physik. Zweite Mitteilung" (Göthingen Nachrichten, 1916) [Translated into Russian in *Osnovaniya Geometrii* (Foundations of Geometry) (Moscow – Leningrad: Gostekhteorizdat, 1948)].

[8] See also new Russian translation by Yu A Danilov and D V Zharikov, in: D Hilbert *Selected Works* Vol. 2 *Analysis, Physics, Problems, Personalia* (Ed. A N Parshin) (Moscow: Factorial, 1998).

[9] According to recent (1921) data up to $1/(3 \times 10^7)$. For data from 1999 see Physics News on the Internet at p. 1284 of this issue of our magazine.

Let us imagine one coordinate system K which rests, and another coordinate system K' which is in a state of uniformly accelerated and rectilinear motion in respect to the first system. A material point moving along a straight line in K moves along a parabola in K'. Taking the direction of motion of K' or the parabola axis for x-axis, one gets in system K'

$$\frac{\mathrm{d}^2 x'}{\mathrm{d}t^2} = g = \mathrm{const}\,.$$

If m is the mass of the point, then each equation written as $m\,\mathrm{d}^2x'/\mathrm{d}t^2 = mg$ can be considered as symbolizing the equality between the product of inert mass by acceleration and the force mg. On the basis of the principle of equivalence, the force mg can be considered as gravitational, with m in this case meaning not the inert mass but an active one which exerts the force mg in the same way as the force mg is exerted by an active mass m of some heavy body on the Earth's surface (in the latter case mg means mM/r^2, where m is the mass of the heavy body, M is the mass of the Earth, and r is the radius of the Earth; it is clear that in this case m exerts force).

Thus, the observer inside the K' system may take no notice of its accelerated and rectilinear motion, if one accepts that such accelerated motion is equivalent to the presence of a gravitating field and that the observer explains all the phenomena which take place around him just on the basis of this gravity. The principle of equivalence enables us to consider the centrifugal force as one which according to its nature coincides with gravitational force and essentially is not different from it; at last, the same can be said about all forces which arise kinematically in a coordinate system linked with a moving body.

There are masses in nature which exert around them the so-called gravitating field; if we take some coordinate system K^*, then the character of the gravitating field will depend on what coordinate system we chose; in another coordinate system $K^{*\prime}$ moving in respect to the first one there will be another gravitating field. Moving together with $K^{*\prime}$, we can ascribe everything which happens in $K^{*\prime}$ not to the motion of $K^{*\prime}$ in respect to K^* but to the gravitating field which exists in $K^{*\prime}$ and differs from the field in K^*.

But the transition from one coordinate system to another arbitrarily chosen coordinate system implies the change of the form of those functions a_{ik} which determine the properties of the geometry of physical space; if in one coordinate system

$$\mathrm{d}s^2 = \sum_{ik} a_{ik}\,\mathrm{d}x_i\,\mathrm{d}x_k\,, \qquad i,k = 1,2,3,4\,,$$

then in $K^{*\prime}$ we obtain

$$\mathrm{d}s^2 = \sum_{ik} a'_{ik}\,\mathrm{d}x'_i\,\mathrm{d}x'_k\,;$$

and evidently $a'_{ik} \neq a_{ik}$, because the dependence between x_i and x'_i is arbitrary.

It should be admitted that the transition from one coordinate system to another changes not only the gravitating field but the geometry of physical space as well, and this indicates that there should be a link between the gravitating field and a_{ik}, i.e. geometry.

Relying on this basis, Einstein calls the quantities a_{ik} gravitational potentials and denotes them g_{ik} by analogy with Earth's acceleration, but this term does not contain anything except for the indicated parallelism between geometry and gravity.

The second Einstein fundamental proposition. Thus, considering the Mach paradox, Einstein arrives at a conclusion *on the admissibility* not only of transition from one uniformly and rectilinearly moving coordinate system to another coordinate system of the same kind, but to ***all coordinate transformations in general*** (because the latter include motion, this means that the new coordinates x'_i, $i = 1,2,3,4$ can be arbitrary functions of four coordinates x_i, $i = 1,2,3,4$).

The third Einstein fundamental proposition. Considering the equivalence principle, Einstein comes to the conclusion that ***the element of arc, underlying the properties of physical space***, i.e.

$$\mathrm{d}s^2 = \sum_{ik} g_{ik}\,\mathrm{d}x_i\,\mathrm{d}x_k\,, \qquad i,k = 1,2,3,4\,,$$

includes 10 functions g_{ik} which determine not only the form of the geometry but the gravitating field in a given coordinate system as well.

The fourth Einstein fundamental proposition. To erect mechanics and physics on this basis, it is necessary to make one essential reservation. If the choice of coordinate system is arbitrary, then how can one describe nature using it? How one can achieve results which do not depend on the established arbitrariness? Because it is quite clear that the laws of nature do not depend on it. The answer inevitably comes to mind: ***since the laws of nature do not depend on our arbitrary choice, they must be independent of the coordinate system chosen by us. Speaking mathematically, the laws of nature should be invariant in respect to any coordinate transformations.*** Einstein genius managed to find and to state the laws of mechanics and physics in just that form invariant and independent of a chosen coordinate system. Now we proceed to the exposition of the fundamental equations of mechanics and physics. All said above only explains the path travelled by Einstein but cannot serve as proof of the validity of his propositions, though principally his statements have an ***obvious*** advantage over corresponding statements of Newtonian mechanics.

The fundamental Einstein equations. Further we shall not follow the path travelled by Einstein but instead shall take after Hilbert who gave a clear and simple exposition of Einstein's theory. In his original work Einstein takes as a starting point the Poisson equation $\Delta\varphi = 4\pi\varrho$, where φ is the normal gravitational potential, and ϱ is the density of matter. Generalization of this equation by means of introduction of 10 potentials g_{ik} instead of φ, and 10 other quantities, determining the state of matter, instead of ϱ gives Einstein the possibility to derive the equations wanted and to show their validity. But the process of generalization is not so simple and so unambiguous in order that one could easily estimate all the significance of the results achieved in this way.

Let us, following Hilbert [10], suppose that all the events which happen in nature depend on some 'world' function H; this function H depends on three coordinates x_1, x_2, x_3 which are purely spatial ones, the fourth coordinate x_4 in the given coordinate system means time. The function H does not depend on the coordinate system chosen by us, and it may be shown that for this reason it does not depend explicitly on x_1, x_2, x_3, x_4; it depends on them only via following quantities.

[10] Einstein, Lorentz and others use the same way of exposition. We follow Hilbert who was the first author who applied this method.

(1) 10 functions g_{ik} and their derivatives over x_i; we could also suppose that H depends on derivatives of any order, but by analogy to Poisson equation we suppose that H depends only on g_{ik} and their first and second order derivatives. We suppose also that these g_{ik} as well as their derivatives are everywhere single-valued and continuous.

(2) Those parameters which determine the state of matter. For example, such parameters are the density of matter, the density of electricity, and electric potentials (vector-potential and scalar potential); if the theory of matter cannot get by with only these parameters, then one has to incorporate the other necessary parameters as well; if one holds the viewpoint of the electromagnetic theory of matter, then it would be sufficient to use as the parameters only the electric density and vector and scalar potentials. At last, sharing Mie's viewpoint, to create the theory of matter it would be sufficient to know the vector-potential and scalar potential, and because the first potential has 3 components, then one should know at all four parameters q_1, q_2, q_3, q_4 as functions of x_1, x_2, x_3, x_4. Taking Mie theory as a basis, Hilbert supposes that H depends on q_1, q_2, q_3, q_4 and their first derivatives over x_i. But for many problems solved by Einstein's theory this supposition is not essential at all.

Thus, let us suppose that there is the 'world' function

$$H = H\left(g_{ik}, \frac{\partial g_{ik}}{\partial x_l}, \frac{\partial g_{ik}}{\partial x_l \partial x_m}, q_i, \frac{\partial q_i}{\partial x_k}\right),$$

where $i, k, l, m = 1, 2, 3, 4$.

Let us consider the integral

$$J = \int H\sqrt{g}\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3\, \mathrm{d}x_4\,, \tag{1}$$

where $\mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3\, \mathrm{d}x_4$ is an element of volume, g is the determinant formed from all g_{ik}, and H, by definition, is an invariant. One can show that $\sqrt{g}\, \mathrm{d}x_1\, \mathrm{d}x_2\, \mathrm{d}x_3\, \mathrm{d}x_4$ is also an invariant, i.e. does not depend on the coordinate system. Obviously J as well as any variation of this integral, is also an invariant.

All phenomena which take place in nature occur in such a way that the variation of this integral δJ is equal to zero:

$$\delta J = 0\,. \tag{2}$$

This is the main law of Einsteinian physics. It should substitute all other laws of physics: Newton's law of gravitation, Maxwell equations, the law of mass interaction and so on.

In order that this law could be of practical significance, one should, of course, know the expression for the function H. Let us suppose that we know it. The expression for H includes 10 unknown functions g_{ik} and 4 unknown functions q_i, but from the condition $\delta J = 0$ follow 14 differential equations, the first 10 of them can be obtained by means of variation of the functions g_{ik}; for brevity we shall denote them

$$G_{ik} = 0\,^{11}, \quad i, k = 1, 2, 3, 4\,; \tag{3}$$

the last 4 equations can be obtained by means of variation of the functions q_i; we shall denote them

$$Q_i = 0\,. \tag{4}$$

[11] We have 10 equations instead of 16 because g_{ik} as well as G_{ik} are symmetric in respect to indices i and k.

The set of equations (3) and (4) enables us to determine g_{ik} and q_i in a given coordinate system.

Equations (3) and (4) derived from the invariant $\delta J = 0$ are invariants themselves (we shall not explain the more exact sense of this statement in this paper due to the lack of space) and do not depend on the coordinate system chosen by us. The arbitrariness of the chosen coordinate system manifests itself here in the fact that these 14 equations are not independent of each other but linked with 4 identities. This means that 4 of the 14 functions g_{ik} and q_i can be chosen arbitrarily and are not determined from equations (3) and (4). The chosen coordinate system is fixed by arbitrary values of 4 of 14 functions.

It may seem at first sight that the determination of the form of the 'world' function H should meet insurmountable difficulties. Meanwhile, the choice of the function H for a very large class of phenomena is almost unambiguous. In fact, let us consider the case (we consider such a case as possible, and the results of the theory imply that this situation really occurs) when the parameters q_1, q_2, q_3, q_4 are small quantities, and instead of them we introduce the quantities εq_1, εq_2, εq_3, εq_4, where ε is some small number and q_1, q_2, q_3, q_4 have finite values. Let us expand H over increasing powers of ε; we arrive at

$$H = K' + \varepsilon L + \varepsilon^2 M + \ldots$$

Let us consider only the first terms of this expansion K' and εL. Then K' depends only on g_{ik} and the first and the second derivatives of these functions over x_i; L depends on g_{ik} and their derivatives, q_i and their derivatives. H is an invariant; $K, L, M, \ldots$ should also be invariants. It turns out that ***there exists only one invariant K' depending on g_{ik}, their first and second derivatives and containing the second derivatives only linearly***; this fact is remarkable. This only invariant is the so-called Rieman curvature of 4-dimensional space. Let us denote it K. It is evident that K' may be equal to K or $K + \lambda$, where λ is some constant number which does not depend on x_i. We shall put $\lambda = 0$; later Einstein and Weyl revealed in their works what great significance the constant λ has; in this paper due to the lack of space we have to abandon this problem and put

$$K' = K\,.$$

Let $D_{\mu\nu}$ be a minor of the determinant g formed from g_{ik}, corresponding to the term $g_{\mu\nu}$ of the determinant; let us denote $D_{\mu\nu}/g$ as $g^{\mu\nu}$ and introduce following notation as well.

Let

$$\begin{bmatrix} ik \\ m \end{bmatrix} = \frac{1}{2}(g_{imk} + g_{mki} - g_{ikm})\,,$$

and let

$$\begin{Bmatrix} ik \\ m \end{Bmatrix} = \sum_m g^{nm} \begin{bmatrix} ik \\ m \end{bmatrix}, \quad i, k, m, n = 1, 2, 3, 4\,.$$

One can show that

$$K = -\frac{1}{2}\sum_{ik} g^{ik} K_{ik}\,,$$

where

$$K_{ik} = \sum_l \frac{\partial}{\partial x_i}\begin{Bmatrix} kl \\ l \end{Bmatrix} - \frac{\partial}{\partial x_l}\begin{Bmatrix} ik \\ l \end{Bmatrix} + \sum_{lm}\begin{Bmatrix} kl \\ m \end{Bmatrix}\begin{Bmatrix} mi \\ l \end{Bmatrix} - \begin{Bmatrix} ik \\ m \end{Bmatrix}\begin{Bmatrix} ml \\ l \end{Bmatrix}.$$

K_{ik} is called the Rieman tensor of curvature: the derivation of this formula can be found in Bianchi's differential geometry. As one can see, the expression for K in the general case is very complicated, but for the solution of some special problems it becomes rather simplified.

The expression for L requires especial consideration. However, following Mie theory, to find L is not difficult. K Schwarzschield has shown that the Maxwell equations could be derived from a kind of Hamiltonian principle. Let us denote the components of the vector-potential $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ as q_1, q_2, q_3, and scalar potential φ as q_4; let the vector r_1, r_2, r_3 mean the electric transfer current ϱv_1, ϱv_2, ϱv_3, where ϱ is the electric density, $\mathbf{v}$ is the normal velocity, and r_4 is equal to ϱ; at last, let

$$M_{ik} = \frac{\partial q_k}{\partial x_i} - \frac{\partial q_i}{\partial x_k}\,. \tag{a}$$

Let us consider the integral

$$L' = \int\left(\sum_{ik} M_{ik}^2 - \sum_i r_i q_i\right) \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_4$$

and put $\delta L' = 0$.

Variation of this integral gives us Maxwell equations

$$\sum_i \frac{\partial M_{ik}}{\partial x_i} = -r_k\,, \tag{b}$$

$$\frac{\partial M_{ik}}{\partial x_l} + \frac{\partial M_{kl}}{\partial x_i} + \frac{\partial M_{li}}{\partial x_k} = 0\,. \tag{c}$$

(In the usual notation instead of (a) one writes

$$\mathbf{E}_1 = -\frac{\partial \varphi}{\partial x_1} - \frac{\partial \mathbf{A}_1}{\partial t} = \frac{\partial q_4}{\partial x_1} - \frac{\partial q_1}{\partial x_4} = M_{14}\,,$$

$$\mathbf{H}_1 = \frac{\partial \mathbf{A}_3}{\partial x_2} - \frac{\partial \mathbf{A}_2}{\partial x_3} = M_{32} \quad \text{and so on}\,;$$

instead of (b) and (c) one writes

$$\operatorname{curl}\mathbf{H} = \frac{\partial \mathbf{E}}{\partial t} + \varrho\mathbf{v}\,,$$

$$\operatorname{curl}\mathbf{E} = -\frac{\partial \mathbf{H}}{\partial t}\,,$$

$$\operatorname{div}\mathbf{E} = \varrho\,,$$

$$\operatorname{div}\mathbf{H} = 0\,.)$$

In the space free of electricity the second term in L_1 drops, and we obtain the Maxwell equations in vacuum.

Mie in his theory also considers such a function L but substitutes some function f of q_i instead of the second term $\sum_i r_i q_i$; as a result, the electric density turns out to be function of the potential q_i. But Mie created his theory not for the general principle of relativity; in his work he was guided by the first 'special' principle of relativity, which is why his L' cannot be directly transferred into the expression for H in the case of the Hilbert 'world' function. To employ the Mie function in the general principle of relativity, the latter should be properly generalized, and in the theory of invariants one can easily prove that such a generalized and, therefore, invariant in respect to any transformation expression will be

$$L = \int\left[\sum_{iklm} M_{ik} M_{lm}\, g^{il} g^{km} - f\left(\sum_{ik} g^{ik} q_i\, q_k\right)\right] \times \sqrt{g}\;\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_4\,. \tag{5}$$

This very expression Hilbert inserts as the second term in the expression of the function H. The dimension of (5) differs from the dimension of K; in order both the dimensions would coincide, L should be multiplied by some numerical coefficient ε. It turns out that $\varepsilon = 8\pi k/c^2$, where k is a gravitational constant, c is the speed of light, i.e. $\varepsilon = 1.87 \times 10^{-27}$ is an extremely small quantity which corresponds to our expansion of the 'world' function H in infinite series [12].

It is remarkable that the number of such invariants L, which can be obtained using q_i and their first derivatives, is also strongly restricted. Mie counts four invariants in total, but chooses from them that invariant which gives him the Maxwell equations at once.

Not holding to the Mie electric theory of matter, one can give L the other form. For some problems, for example, for astronomical ones, de Sitter, Einstein and others do just so; but, as we shall show later, the form of the function L does not play any role for the solution of the most simple and interesting astronomical problems.

Thus, let us put

$$H = K + \varepsilon L\,,$$

where K is the curvature of 4-dimensional space, and L is given by (5).

Examples. Now we can proceed to the solution of some individual problems which should show what can be given by Einstein's theory and how it leads to the solution of mechanical and physical problems.

The 1st example. Let us suppose that space is free of matter, then $L = 0$, and we only have

$$J = \int K\sqrt{g}\;\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\mathrm{d}x_4\,.$$

In this case $\delta J = 0$ implies 10 equations

$$G_{ik} = 0\,.$$

If one assumes (and this should be done in according with the essence of the theory) that g_{ik} are continuous and single-valued functions, then the solution of these differential equations will be

$$g_{ik} = 0 \quad \text{for} \quad i \neq k\,, \quad \text{and}$$

$$g_{ii} = 1 \quad \text{for} \quad i = 1, 2, 3\,, \quad \text{and}$$

$$g_{44} = -1$$

(the value -1 and not $+1$ for g_{44} stems from those requirements which should be satisfied by g_{ik} in order that

[12] To prove that $\varepsilon = 8\pi k/c^2$, one can use some very simple examples; we shall omit the proof due to a lack of space.

x_4 mean time). Thus we get

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - dx_4^2\,,$$

i.e. the same expression for ds^2 which we have in Einstein's 'special' principle of relativity. It does not include the speed of light, because we put it equal to unity, which evidently affects only the choice of the measurement unit for x_4, i.e. time.

Therefore, in the absence of matter we have the usual expression for ds^2*, i.e. Euclidean geometry in a 3-dimensionally extent space*[13].

The 2nd example. Let us suppose that we consider a space inside some very small sphere circumscribed around a point x_1, x_2, x_3, x_4; if the radius of this sphere is small enough, then inside the sphere the quantities g_{ik} can be considered as constant. In this case it is easily shown that the expression for ds^2 can always be transformed into

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - dx_4^2\,.$$

To this end, it is the sufficient transformation. From this fact one can draw the conclusion that in the infinitesimal there is always valid 'small' principle of relativity. In this case the 'world' function transforms into

$$H = \varepsilon \int L\, dx_1\, dx_2\, dx_3\, dx_4$$

and, carrying out its variation, we come to the usual Maxwell equations, because all the g_{ik} are equal to either unity (for $i = k$) or zero (for $i \neq k$): if we put speed of light not equal to 1, then

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - g_{44}\, dx_4^2\,;$$

obviously, $g_{44} = c^2$; the same g_{44}, as one can see from expressions (5) for L, enters the Maxwell equations; now it is clear why in the special principle of relativity a 'physical' quantity enters the geometrical expression for ds^2; the statement that the speed of light c is a constant quantity is also clear. It is constant because we have the right to consider the g_{ik} independently of coordinates x_i.

The 3rd example. Let us consider the so-called one-body problem, i.e. the gravitational field exerted by one gravitating mass. Let this mass be located at the origin of the coordinate system, and let it have a spherical shape. The gravitational field it exerts should be spherically symmetric but only if one supposes that the sphere is at rest and everything is in a stationary state, i.e. all the g_{ik} do not depend on t. Let us introduce the conditions of spherical symmetry into the expression for ds^2. According to Schwarzschield, if one introduces the spherical coordinates

$$x_1 = r\cos\vartheta\,,$$
$$x_2 = r\sin\vartheta\cos\varphi\,,$$
$$x_3 = r\sin\vartheta\sin\varphi$$

and puts

$$x_4 = t\,,$$

then the most general expression for ds^2 will be[14]

$$ds^2 = F(r)\, dr^2 + G(r)(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) + H(r)\, dt^2\,.$$

But instead of r we can take

$$r' = \sqrt{G(r)}\,,$$

then

$$ds^2 = M(r)\, dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) - W(r)\, dt^2$$

(we drop the prime over r).

Two arbitrary functions of r, i.e. $M(r)$ and $W(r)$, should be determined from variation of the integral J. To find this variation, we should know not only L but the function L as well, but we can proceed here in the same way as in the theory of the potential when solving the Poisson equation $\Delta\Psi = 4\pi\varrho$; instead of determining the continuous and single-valued functions Ψ satisfying this equation, one may consider the equation $\Delta\Psi = 0$, find its solutions and suppose that the singular points are the points of matter concentration ϱ; here we proceed in the same way. We drop the function L but, solving the remaining equations, allow solutions with singular points and suppose that mass is concentrated at these points.

Consequently, we have to solve the problem

$$\delta \int K\sqrt{g}\, dr\, d\vartheta\, d\varphi\, dt = 0\,.$$

To this end we should evaluate the curvature K by starting from those expressions for g_{ik} which enter the expression for ds^2. These calculations are fairly long and in the end bring us to the following expression for $K\sqrt{g}$:

$$K\sqrt{g} = \left\{ \left(\frac{r^2 W'}{\sqrt{MW}}\right) - 2\,\frac{rM'\sqrt{W}}{M^{3/2}} - 2\sqrt{MW} + 2\sqrt{\frac{W}{M}} \right\} \sin\vartheta\,;$$

let us introduce the functions $m(r)$ and $w(r)$ instead of M and W in such a way that

$$M = \frac{r}{r-m} \qquad \text{and} \qquad W = w^2\,\frac{r-m}{r}\,.$$

This gives

$$K\sqrt{g} = \int\!\left\{ \left(\frac{rW'}{\sqrt{MW}}\right)' - 2m'w \right\} \sin\vartheta\,.$$

The prime $'$ here means differentiation over r. Carrying out all possible integrations, we obtain at last

$$\delta \int K\sqrt{g}\, dr\, d\vartheta\, d\varphi\, dt = -\delta \int 2m'w\, dr = 0\,,$$

and this gives two differential equations

$$m' = 0 \qquad \text{and} \qquad w' = 0\,,$$

[13] In its latest development Einsteinian theory comes to the conclusion that in the absence of matter all $g^{\mu\nu}$ are equal to zero, i.e. there is no physical space at all without matter. In principle this is, of course, the only right conclusion.

[14] See also D Hilbert, *loc.cit.*

i.e. $m = \mathrm{const}$ and $w = \mathrm{const}$. Let us put $m = \alpha$ and $w = 1$; the latter does not restrict our problem, because it is obvious that only a choice of the unit of time is connected with the value of w.

As a result, we arrive at the following expression

$$\mathrm{d}s^2 = \frac{r}{r-\alpha}\,\mathrm{d}r^2 + r^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2) - \frac{r-\alpha}{r}\,\mathrm{d}t^2\,. \quad (6)$$

We see that the solution of our problem brings us to the functions g_{ik} having a singular surface — a sphere of radius α; on this sphere, masses are concentrated, exerting a gravitational field with spherical symmetry. If one puts $\alpha = 0$, i.e. accepts that there is no singular surface, then the functions g_{ik} become continuous and single-valued functions having no singular points, but at the same time taking that very magnitude which they have in Euclidean geometry and which, as we can see, corresponds to the absence of matter.

Thus, the gravitational field is determined, and we now have to consider the laws of motion of material particles in such a field, which do not disturb it. To find such laws let us suppose that the motion of particles proceeds in the same way as in Newtonian mechanics, when no forces exert their action on the particles, i.e. we suppose that ***the particles move along the shortest paths or geodesics***: this means that

$$\delta\int \mathrm{d}s = 0\,,$$

and we have to solve a new variational problem.

Let us consider r, φ, ϑ, t as functions of the same parameter p; our task is to solve the set of differential equations derived from the condition

$$\delta\int\left\{\frac{r}{r-\alpha}\left(\frac{\mathrm{d}r}{\mathrm{d}p}\right)^2 + r^2\left[\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}p}\right)^2 + \sin^2\vartheta\left(\frac{\mathrm{d}\varphi}{\mathrm{d}p}\right)^2\right] - \frac{r-\alpha}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}p}\right)^2\right\}^{1/2}\mathrm{d}p = 0\,.$$

It is easy to show that the geodesics obtained in this way will be planar. But then we can restrict ourselves to those shortest curves which lay in the plane of equator, and put $\vartheta = \pi/2$. The former expression can be rewritten as

$$\delta\int\sqrt{\frac{r}{r-\alpha}\left(\frac{\mathrm{d}r}{\mathrm{d}p}\right)^2 + r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}p}\right)^2 - \frac{r-\alpha}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}p}\right)^2}\,\mathrm{d}p = 0\,.$$

From here follow three second-order differential equations

$$\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{2r}{r-\alpha}\frac{\mathrm{d}r}{\mathrm{d}p}\right) + \frac{\alpha}{(r-\alpha)^2}\left(\frac{\mathrm{d}r}{\mathrm{d}p}\right)^2 - 2r\left(\frac{\mathrm{d}\varphi}{\mathrm{d}p}\right) + \frac{\alpha}{r^2}\left(\frac{\mathrm{d}t}{\mathrm{d}p}\right)^2 = 0\,,$$

$$\frac{\mathrm{d}}{\mathrm{d}p}\,r^2\,\frac{\mathrm{d}\varphi}{\mathrm{d}p} = 0\,, \qquad \frac{\mathrm{d}}{\mathrm{d}p}\,\frac{r-\alpha}{r^2}\,\frac{\mathrm{d}t}{\mathrm{d}p} = 0\,. \quad (7)$$

Three first integrals of these equations are

$$\frac{r}{r-\alpha}\left(\frac{\mathrm{d}r}{\mathrm{d}p}\right)^2 + r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}p}\right)^2 - \frac{r-\alpha}{r}\left(\frac{\mathrm{d}t}{\mathrm{d}p}\right)^2 = A\,,$$

$$r^2\,\frac{\mathrm{d}\varphi}{\mathrm{d}p} = B\,, \qquad \frac{r-\alpha}{r}\,\frac{\mathrm{d}t}{\mathrm{d}p} = C\,, \quad (7')$$

where A, B, C are constants of integration. The value of the constant C determines only the choice of the units for the parameter, so we put $C = 1$.

Eliminating p and t from these equations, one obtain the equation for the trajectory of motion; after obtaining, let us make the substitution $1/r = \rho$ and then, at last, get

$$\left(\frac{\mathrm{d}\rho}{\mathrm{d}\varphi}\right)^2 = \frac{1+A}{B^2} - \frac{A\alpha}{B^2}\,\rho - \rho^2 + \alpha\rho^3\,. \quad (8)$$

This relation strongly resembles the Kepler equation for planetary motion. The latter is derived and written in the same way.

The law of conservation of energy gives

$$\frac{m}{2}\left[\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2\right] - k\,\frac{Mm}{r} = a\,,$$

where a is the energy; m is the mass of the planet, which will further be taken equal to 1; M is the mass of the Sun, and k is the gravitational constant.

The law of conservation of area gives

$$r^2\,\frac{\mathrm{d}\varphi}{\mathrm{d}t} = b\,,$$

eliminating t and putting $\rho = 1/r$, we obtain

$$\left(\frac{\mathrm{d}\rho}{\mathrm{d}\varphi}\right)^2 = \frac{a}{b^2} - \frac{kM}{b^2}\,\rho - \rho^2\,.$$

Let us insert new constants a and b into (8) instead of constants A and B in such a way that

$$\frac{A\alpha}{B^2} = \frac{kM}{b^2}\,, \qquad 1 + A = \frac{\alpha A}{kM}\,,$$

then Eqn (8) transforms into

$$\left(\frac{\mathrm{d}\rho}{\mathrm{d}\varphi}\right)^2 = \frac{a}{b^2} + \frac{kM}{b^2}\,\rho - \rho^2 - \alpha\rho^3\,.$$

It is evident that the last expression in the limit, when $\lim \alpha = 0$, becomes the Kepler equation of planetary motion. If α is very small, then the last term of the equation can be dropped if ρ cannot become very large, i.e. in the case when the planet does not approach very close to the Sun.

Let us find the physical meaning of the quantity α. To this end, let us consider circular motion. One can show that $r = \mathrm{const}$ is an integral of the differential equations (7) and, therefore, circular motion is possible, but in this case equation (7) gives us

$$r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2 = \frac{\alpha}{2r}$$

in which the unit of time is chosen in such a way that $c = 1$; if $c \neq 1$, then

$$r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2 = \frac{\alpha}{2r}\,c^2\,.$$

But from the Kepler equation it follows that for a circular motion

$$r^2\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2 = k\,\frac{M}{r^2}\,.$$

Comparing the last two formulas, we obtain

$$\frac{\alpha}{2} = \frac{kM}{c^2} (= 1.5 \times 10^5 \text{ cm for the Sun}) .$$

This means that here the constant α plays the role of the mass of the Sun and is measured in cm. For the Sun it is equal to 1.5 km.

Really, for all planets the value of α is very small in comparison with the radius-vector of their orbits, and the Kepler equation should be valid with very high accuracy. Nevertheless, the additional term $\alpha\rho^3$, not entering the classical equation, can be significant in some special cases. Because the value of α for all known planets really is small in comparison with the radius-vectors of their orbits, to take the influence of α into account, we can when solving differential equations (8) use the expansion in powers of α; let e_1, e_2, e_3 be the roots of the equation

$$f(\rho) = \frac{\alpha}{b^2} + \frac{kM}{b^2}\,\rho - \rho^2 + \alpha\rho^3 = 0 .$$

It is obvious that

$$e_1 + e_2 + e_3 = +\frac{1}{\alpha}$$

and

$$f(\rho) = (\rho - e_1)(t_2 - \rho)\left[1 - \alpha(\rho + e_1 + e_2)\right] .$$

The equation of motion is

$$\mathrm{d}\varphi = \frac{\mathrm{d}\rho}{\sqrt{(\rho - e_1)(e_2 - \rho)\left[1 - \alpha(\rho + e_1 + e_2)\right]}} . \tag{9}$$

Evidently, the motion takes place between $\rho = e_1$ and $\rho = e_2$. Expanding in a series over powers of α, we obtain [15]

$$\mathrm{d}\varphi = \frac{\mathrm{d}\rho}{\sqrt{(\rho - e_1)(e_2 - \rho)}}\left[1 + \frac{\alpha}{2}(e_1 + e_2) + \frac{\alpha}{2}\,\rho\right] ,$$

and integrating results in

$$\varphi - \varphi_0 = -\frac{\alpha}{2}\sqrt{(\rho - e_1)(e_2 - \rho)} + \left[1 + \frac{3}{4}\,\alpha(e_1 + e_2)\right] \arcsin\frac{(e_1 + e_2)/2 - \rho}{(e_1 - e_2)/2} .$$

This formula enables us to evaluate the angle Φ between the radius-vectors of the points of the greatest and the least distance from the Sun, i.e. between $\rho = e_1$ and $\rho = e_2$. It is obvious that

$$\Phi = \pi\left[1 + \frac{3}{4}\,\alpha(e_1 + e_2)\right] .$$

Returning to the point of the greatest distance (perihelion), the planet turns at the angle

$$2\Phi = 2\pi\left[1 + \frac{3}{4}\,\alpha(e_1 + e_2)\right] .$$

For Keplerian motion we have the corresponding angle $2\Phi k = 2\pi$; thus, we see that according to Einstein's theory, during one revolution of the planet around the Sun the perihelion of the orbit shifts by the angle

$$\omega = \frac{3}{2}\,\alpha(e_1 + e_2)\pi .$$

Let T be the period of the planet's revolution, a be the major semiaxis of the orbit, and ε be the eccentricity of the orbit. Then

$$\alpha = \frac{kM}{c^2} = \frac{(2\pi)^2 a^2}{T^2 c^2} ,$$

$$e_1 + e_2 = \frac{2}{a(1 - \varepsilon^2)} .$$

After substitution we get

$$\omega = 24\pi^3\,\frac{a^2}{T^2 c^2 (1 - \varepsilon^2)} ;$$

this quantity is very small; for the planet Mercury over one hundred years, i.e. for $\Omega = (100/T')\omega$, where T' is the period of revolution for Mercury (expressed in terrestrial years) we get

$$\Omega = 43'' .$$

This value agrees perfectly with experiment and cannot be explained by any other theory without introducing new hypothesis ad hoc**!** [16]

For other planets the quantity Ω is considerably less, and an experimental test cannot be of such crucial significance as for Mercury.

Let us consider the rectilinear motion of a mass point falling directly on the Sun; it can be shown that geodesics corresponding to such motion are possible. Then $\varphi = \text{const}$, and the dependence of r on t is determined from the equation

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} = \frac{3\alpha}{2r(r - \alpha)}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 - \frac{\alpha(r - \alpha)}{2r^3} .$$

Here the usual speed of light is taken as unity; we see that if

$$\left|\frac{\mathrm{d}r}{\mathrm{d}t}\right| < \frac{1}{\sqrt{3}}\,\frac{r - \alpha}{r} = \frac{c_r}{\sqrt{3}} ,$$

then the acceleration will be positive; but if

$$\frac{\mathrm{d}r}{\mathrm{d}t} > \frac{c_r}{\sqrt{3}} ,$$

then the acceleration will be negative; the quantity c_r, as will be shown later, coincides with the speed of light at the point r.

The mass of the moving planet is assumed in our calculations to be equal to unity. Therefore, from the viewpoint of ordinary mechanics we can consider the expression for $\mathrm{d}^2 r/\mathrm{d}t^2$ as the expression for the force acting on a unit mass. We see that Einstein's theory manages without the notion of a force, but a question can arise of whether the

[15] We drop the terms containing α^2.

[16] Really, the observed motion of Mercury's perihelion is somewhat more, but the difference between the observed value and the angle $43''$ is explained by the perturbing influence of other planets; this remainder amounting to $43''$ cannot be explained in the framework of classical theory.

corresponding change of Newton's law can bring us to the same results as Einstein's theory. The answer is negative. Really, in the case of rectilinear motion the relation for a Newtonian force is

$$F_{\mathrm{d}} = -\frac{\alpha}{2r^2} + \frac{\alpha^2}{2r^3} + \frac{3\alpha}{2r(r-\alpha)}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 ;$$

in the limit, for very small α, this relation becomes the classical expression

$$F = -\frac{kM}{r^2} .$$

But, as was shown, for circular motion according to Einstein's theory the force should be

$$F_c = -\frac{\alpha}{2r^2} .$$

Evidently, F_c and F_d cannot be particular cases of the same general law. If one drops the terms containing the radial velocity, i.e. $\mathrm{d}r/\mathrm{d}t$, becoming zero for circular motion, in the expression for F_d, then the rest will be the expression for F_c. According to Einstein's theory, the expression for the force (if one insists on the introduction of the term 'force' and gives it the meaning of the quantity equal to the production of mass by acceleration) becomes dependent on the trajectory of the mass point, i.e. it does not have the meaning of the universal law in the same sense as the Newtonian law of gravitation. It is self-evident that in the limit, i.e. for very small α, F_c and F_d coincide and give F.

The 4th example. Let us turn now to the consideration of light motion. Light, like a mass point, moves along geodesics, but in contrast to a mass point and in the same way as in the special principle of relativity, the length of these geodesic lines is equal to zero, and we have for them

$$\mathrm{d}s^2 = 0 .$$

According to this, we have to put $A = 0$ in the integrals of equations (7), and the trajectories of the light beams will be curves determined by integration of the expression

$$\left(\frac{\mathrm{d}\rho}{\mathrm{d}\varphi}\right)^2 = \frac{1}{B^2} - \rho^2 - \alpha\rho^3 ;$$

the latter in the limit for $\lim a = 0$ can be integrated very easily, and we get

$$B\rho = \sin(\varphi - \varphi_0) ,$$

where φ_0 is the constant of integration, i.e. simply a straight line

$$r = \frac{B}{\sin(\varphi - \varphi_0)} .$$

The quantity B here means the shortest distance of the ray from the Sun.

Let us consider now not the limiting case $\alpha = 0$, but suppose only that α is enough small in comparison with the closest trajectory to the Sun. Let e_1, e_2, e_3 be the roots of the equation

$$\frac{1}{B^2} - \rho^2 + \alpha\rho^3 = 0$$

and let e_1 and e_2 become in the limit when $\lim \alpha = 0$ the roots of the limiting equation

$$\frac{1}{B^2} - \rho^2 = 0 ,$$

i.e. let

$$\lim e_1 = \frac{1}{B} \quad \text{and} \quad \lim e_2 = -\frac{1}{B} .$$

The equation

$$\frac{\mathrm{d}\rho}{\sqrt{1/B^2 - \rho^2 + \alpha\rho^3}} = \mathrm{d}\varphi \tag{10}$$

can be approximately integrated quite in the same way as equation (9) on page 1243. Obviously, we obtain

$$\varphi - \varphi_0 = -\frac{\alpha}{2}\sqrt{(\rho - e_1)(e_2 - \rho)} + \left[1 + \frac{3}{4}\,\alpha(e_1 + e_2)\right] \arcsin\frac{(e_1 + e_2)/2 - \rho}{(e_1 - e_2)/2} , \tag{9$'$}$$

approximate values for e_1 and e_2 can easily be evaluated; we get

$$e_1 = \frac{1}{B} - \frac{1}{2}\frac{\alpha}{B^2} , \qquad e_2 = -\frac{1}{B} - \frac{1}{2}\frac{\alpha}{B^2} .$$

Writing (9$'$) as

$$r = \frac{2}{e_1 + e_2} \times \left\{1 - \frac{e_2 - e_1}{e_2 + e_1}\sin\left[\varphi - \varphi_0 + \frac{\alpha}{2}\sqrt{(\rho - e_1)(e_2 - \rho)}\right]\right\}^{-1} ,$$

one can see at once that we are dealing with a curve resembling a hyperbola with eccentricity

$$\varepsilon = \frac{e_2 - e_1}{e_2 + e_1} = \frac{2B}{\alpha} ;$$

B means approximately the shortest distance of the trajectory from the Sun, i.e. is supposed to be very large in comparison with α; thus, this hyperbola has very a large eccentricity and differs very little from a straight line.

For the asymptotes of this hyperbola $r = \infty$ and $\rho = 0$; therefore, φ can be determined from the condition

$$1 - \frac{e_2 - e_1}{e_2 + e_1}\sin\left(\varphi - \varphi_0 + \frac{\alpha}{2}\sqrt{-e_1 e_2}\right) = 0 .$$

Let us substitute here the values of e_1 and e_2 and choose the arbitrary constant $\varphi_0 = (\alpha/2)\sqrt{-e_1 e_2}$; the angle φ, which is measured from this arbitrary direction, will be determined from the condition

$$\sin\varphi = \sin(\pi - \varphi) = \frac{e_2 + e_1}{e_2 - e_1} = \frac{\alpha}{2B} .$$

Consequently, the angles which the asymptotes make with the direction φ_0, are very small and equal to

$$\varphi = \pm\frac{\alpha}{2B} ,$$

the angle between them being

$$\Psi = \frac{\alpha}{B}.$$

If a ray of light propagates along such a hyperbola, then the Sun will be located at its focus. For parts of the curve sufficiently distant from the Sun, the motion along the hyperbola can be identified with the motion along the asymptote. Thus, we come to conclusion that ***a ray of light passing near the Sun is deflected through an angle***

$$\Psi = \frac{\alpha}{B} = \frac{kM}{c^2 B},$$

where B is the shortest distance of the ray from the Sun. Einstein evaluated this angle for a ray touching the surface of the Sun and found that $\Psi = 1.7''$; the observations carried out in 1919 by an English expedition to Brazil, brilliantly supported this result predicted by Einstein in advance.

The study of equation (10) determining the motion of light presents much of interest; due to lack of space we cannot describe it here; let us indicate only some peculiar cases. If a ray of light comes close enough to the surface $r = 3\alpha/2$, then it winds around it and cannot deviate from it any more . No ray can penetrate through the surface $r = \alpha$. If the ray goes along a straight line in the direction to the center of the Sun, then its speed is determined by

$$\frac{\mathrm{d}r}{\mathrm{d}t} = c_r = 1 - \frac{\alpha}{r}\,^{17},$$

its acceleration is always negative, and it approaches the surface $r = \alpha$ for an infinitely long time with speed $c_r = 0$.

The planets orbiting the Sun in the circles have a higher speed the closer they are to the Sun; a planet orbiting in a circle of radius $r = 3\alpha/2$ has the speed of light, but this speed of light is not c, but $c/\sqrt{3}$; inside a circle of radius $r = 3\alpha/2$ circular motion is impossible.

We can see that for the Sun $\alpha = 1.5 \times 10^5$ cm; in comparison with the radius of the Sun it is a small magnitude and for this reason the peculiar properties of the surfaces $r = \alpha$ and $r = 3\alpha/2$ do not have practical significance. For a hydrogen molecule $\alpha = 10^{-49}$ approximately.

The 5th example. The special principle of relativity taught us that the time measured by a moving observer and an observer at rest do not coincide. Let x_1, y_1, z_1 be three functions of time t giving us the motion of some point. The element of time, measured by an observer at rest, is $\mathrm{d}t$, meanwhile the element of time $\mathrm{d}\tau$, measured by observer moving together with point, is determined from the relation

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2,$$

where τ is the so-called 'proper' time of the point. In general relativity we also have to distinct between the increment of the fourth coordinate $\mathrm{d}t$ and the 'proper' time $\mathrm{d}\tau$ of some point. The difference between 'general' and 'special' relativity is only that in special relativity for a point at rest $\mathrm{d}t = \mathrm{d}\tau$, i.e. 'proper' time coincides with the increment of the fourth coordinate — time, which is not the case for general relativity. Let us take, for example, the gravitational field discussed in the 3rd example. If a point is at rest, then $\mathrm{d}x$, $\mathrm{d}y$, $\mathrm{d}z$ are equal to zero (or in spherical coordinates $\mathrm{d}r$, $\mathrm{d}\varphi$, $\mathrm{d}\vartheta$ are equal to zero), and the 'proper' time is determined for every point at rest as

$$\mathrm{d}\tau^2 = \left(1 - \frac{a}{r}\right)\mathrm{d}t^2,$$

that is, in other words, the increment of the fourth coordinate — time — is not equal to the increment of the 'proper' time, but depends on r and α, i.e. on the distance from the Sun and mass.

Let us consider some periodical molecular process, say, the radiation process. It is natural to suppose that for a molecule or a particle vibrating in a molecule the period of radiation characteristic for a given molecule and stemming from its intrinsic properties is a period which is determined from its 'proper' time and does not depend on an arbitrarily superimposed coordinate system x, y, z, t and, therefore, on an arbitrarily superimposed coordinate — time t; thus, for the molecule itself the period is the same everywhere; let us suppose that this period is so small that it can be denoted $\mathrm{d}\tau$, and the rate of vibrations is so small that in the relation for $\mathrm{d}\tau$ we can put $\mathrm{d}r = \mathrm{d}\varphi = \mathrm{d}\vartheta = 0$. But our observations are made in a coordinate system chosen by us; the period we measure is not $\mathrm{d}\tau$, but $\mathrm{d}t$; now we write the condition that $\mathrm{d}\tau$ is everywhere the same, taking $\mathrm{d}\tau$ once on the surface of the Sun, i.e. putting $r = d$ (radius of the Sun), and another time putting $r = D$ (semidiameter of the terrestrial orbit). It is obvious that

$$\mathrm{d}\tau^2 = \left(1 - \frac{\alpha}{d}\right)\mathrm{d}t_d^2 = \left(1 - \frac{\alpha}{D}\right)\mathrm{d}t_D^2,$$

where $\mathrm{d}t_d$ and $\mathrm{d}t_D$ are the periods measured, correspondingly, at the Sun and at the Earth. But α/D is a quantity very small in comparison with α/d and can be neglected. On the other hand, if $\mathrm{d}t_d$ and $\mathrm{d}t_D$ are the periods whose inverse quantities are frequencies ν_d and ν_D, then our condition can be written as

$$\nu_d = \nu_D\left(1 - \frac{\alpha}{d}\right)\frac{1}{2} = \nu_D\left(1 - \frac{\alpha}{2d}\right).$$

Let us denote ν_D just by ν, and $\nu_d - \nu_D$ by $\mathrm{d}\nu$; it is evident that

$$\mathrm{d}\nu = -\frac{\alpha}{2d}\,\nu \quad \text{or, for } \nu = \frac{1}{\lambda},$$

$$\mathrm{d}\lambda = +\frac{\alpha}{2d}\,\lambda.$$

The light emission by some luminous gas has the character of the periodical motion we discussed. We see that ***the gravitational potential of the Sun*** $\alpha/2\mathbf{d}$ ***should shift the lines emitted by a gas in the 'red'*** **direction** ($\mathrm{d}\lambda > 0$). A Einstein evaluated this shift, and experiments, in all appearance, have confirmed the result predicted by the theory.

P.S. The article is based on a series of reports made at colloquia headed by Academician P P Lazarev (first Director of the Physics Institute and the first Editor-in-Chief of *Uspekhi Fizicheskikh Nauk* journal) at the Physics Institute of Moscow Science Institute, the first of which took place on 22 October 1918. (*Note by the Editor.*)

[17] One should substitute $A = 0$ in Eqn (7′).