

A new extension of the theory of relativity

By H. WEYL

Translated by D. H. Delphenich

Chapter I. Geometric foundations.

Introduction. – In order to be able to characterize the physical state of a world-location by numbers:

1. The neighborhood of that location must be referred to *coordinates*.
3. Certain *units of measurement* must be established.

Up to now, Einstein’s theory of relativity has addressed only the first point – i.e., the arbitrariness of coordinate systems. However, it is important to ascribe just as prominent of a position to the second point – i.e., the arbitrariness of the units of measurement. We shall speak about that in what follows.

The world is a four-dimensional continuum, and for that reason, can be referred to four coordinates x_0, x_1, x_2, x_3 . The transition to another coordinate system \bar{x}_i will be mediated by continuous transformation formulas:

$$(1) \quad x_i = f_i(\bar{x}_0, \bar{x}_1, \bar{x}_2, \bar{x}_3) \quad (i = 0, 1, 2, 3).$$

There is nothing to distinguish the various possible coordinate systems intrinsically. The relative coordinates dx_i of a point $P' = (x_i + dx_i)$ that is infinitely close to the point $P = (x_i)$ are the components of the infinitesimal displacement $\overline{PP'}$ (i.e., a “line element” at P). Under the transition (1) to another coordinate system \bar{x}_i , they transform linearly:

$$(2) \quad dx_i = \sum_k \alpha_k^i d\bar{x}_k .$$

α_k^i are the values of the derivatives $\partial f_i / \partial \bar{x}_k$ at the point P . The components ξ^i of any *vector* at P transform in the same way. An “axis-cross” at P is linked with a coordinate system in the neighborhood of P that consists of the “unit vectors” ϵ_i with the components $\delta_i^0, \delta_i^1, \delta_i^2, \delta_i^3$:

$$\delta_i^k = \begin{cases} 0 & i \neq k \\ 1 & i = k \end{cases} .$$

It is just that axis-cross that one must use as a basis in order to be able to characterize non-scalar quantities. The transformation formulas:

$$\bar{\mathbf{e}}_i = \sum_k \alpha_i^k \mathbf{e}_k,$$

which are “contragredient” to (2), exist between the unit vectors \mathbf{e}_i , $\bar{\mathbf{e}}_i$ of two coordinate systems at P . In the special theory of relativity, the α_i^k are constants (i.e., independent of position), because the transition functions f_i in (1) are always linear in it; that is not the case in the general theory of relativity.

In order to clarify the dependency of the measurement numbers upon the unit of measurement, we address the geometric example of the *line segment*. **Riemann** ⁽¹⁾ assumed that one could compare the measurements of infinitely-small line segments at the same place, as well as at different places, with each other, and the **Riemannian** geometry that is based upon that assumption lies at the foundations of **Einstein**’s theory of gravitation when it is applied to the four-dimensional continuum. If one establishes a certain line segment as one’s unit of measurement (and naturally, it is the same everywhere) then any line segment will take on a measurement number l that characterizes it completely. However, for a different choice of unit of measurement, one will get a different measurement number \bar{l} that emerges from l by the linear transformation:

$$\bar{l} = a l.$$

In this, a , which is the ratio of the units of measurement, is a universal constant (independent of position and line segment). As one sees, in regard to the question of units of measurement, this viewpoint corresponds precisely to the one that the special theory of relativity assumes in the context of the axis-cross. In place of it, the general theory of relativity postulates only that a is independent of the line segment, but not of position. One must abandon the assumption that “distant comparison” is possible, which is not permissible in a purely “local” geometry, anyway. Only measurements of line segments that are found at the same place can be compared with each other. The gauging of line segments must be performed at each individual world-event, so the problem cannot be handed over to a central gauging authority. However, a principle must enter in place of **Riemannian** distant comparison that allows for the congruent transplantation of the line segments at a point P to the *points that are infinitely-close to P* . With that, I believe, the historical process of the detachment of Euclidian rigidity – i.e., the overcoming of distant geometry – has now terminated. A pure infinitesimal geometry will come about that is, in the same sense, a pure local action theory of physics, just as **Riemannian** geometry is the foundation for the physics that is touched upon in the context of **Einstein**’s general theory of relativity. I shall summarize the main concepts

⁽¹⁾ “Über die Hypothesen, welche der Geometrie zugrunde liegen,” *Mathematische Werke*, 2nd ed., Leipzig, 1892, no. XIII, pp. 272.

and facts of infinitesimal geometry here. A more thorough presentation is contained in the third edition of my book *Raum, Zeit, Materie* (Springer, to be printed) ⁽¹⁾.

Geometry. – A four-dimensional manifold is *affinely-connected* when one can be certain of what vector at a point P' every vector at a point P will go to under parallel translation; in this, P' means an arbitrary point that is infinitely close to P . One demands that there should exist a coordinate system about the point P (which I call “geodetic” at P) such that the components of any vector at P in or will remain unchanged under parallel displacement. If one employs an arbitrary coordinate system x_i , and if $P = (x_i^0)$, $P' = (x_i^0 + dx_i)$ in it, and if an arbitrary vector at P has the components ξ^i , so the vector that emerges from it by parallel translation to P' is $\xi^i + d\xi^i$, then one will have an equation ⁽²⁾:

$$(3) \quad d\xi^i = -d\gamma^i_r \xi^r.$$

The infinitesimal quantities $d\gamma^i_r$, which do not depend upon the vector ξ , are linear differential forms:

$$d\gamma^i_r = \Gamma^i_{rs} dx_s,$$

whose numerical coefficients Γ – viz., the “components of the affine connection” – satisfy the symmetry condition $\Gamma^i_{rs} = \Gamma^i_{sr}$. (3) expresses the idea that the parallel displacement of P to P' maps the set of vectors at P affinely (or linearly) to the set of vectors at P' . If the coordinate system is geodetic at P then all Γ will vanish there. There is no difference between the various points of the manifold in regard to the nature of their affine connection in the vicinity of P .

A *metric* manifold carries a *length determination* at every point P ; i.e., every vector \mathfrak{x} at P determines a *line segment*, and there is a quadratic form \mathfrak{x}^2 (with index of inertia 3), which depends upon the arbitrary vector \mathfrak{x} , in such a way that two vectors \mathfrak{x} and \mathfrak{y} at P will determine the same line segment if and only if $\mathfrak{x}^2 = \mathfrak{y}^2$. The form is established only up to an arbitrary positive proportionality factor in that way. If we choose it in a certain way then the manifold will be *gauged* at P , and we then call $\mathfrak{x}^2 = l$ the *length* of the line segment that is determined by \mathfrak{x} . If one changes the gauge then the same line segment will take on a different length \bar{l} that emerges from l by a linear transformation $\bar{l} = a l$ (a is a positive constant). Relative to a coordinate system, one expresses \mathfrak{x}^2 for the arbitrary vector \mathfrak{x} with the components ξ^i by the formula:

$$\mathfrak{x}^2 = \sum_{ik} g_{ik} \xi^i \xi^k \quad (g_{ki} = g_{ik}).$$

⁽¹⁾ One can also confer the author’s papers in the Sitz. Preuss. Akad. Wiss. (1918), pp. 465, *et seq.* and Math. Zeit. 2 (1918), pp. 384, *et seq.*

⁽²⁾ One always sums over indices that appear twice.

However, not only does a metric manifold carry a length determination at every point, but it is also *metrically connected*. That concept is completely analogous to that of the affine connection; as far as vectors are concerned, they are like the line segments. Any line segment at P then goes to a well-defined line segment at the infinitely-close point P' under congruent transplantation. One must again require that the gauging can be arranged (it will then be called geodetic at P) such that the length of any line segment at P will remain unchanged under congruent transplantation. If the manifold is gauged in some way, and l is the length of a line segment of P , so $l + dl$ is the length of the line segment that arises from it by congruent transplantation to P' , then one will have:

$$(4) \quad dl = -l d\varphi$$

as a result, in which $d\varphi$ does not depend upon the line segment. This equation expresses the idea that any transplantation will define a similarity map of the line segments at P to the line segments at P' . Secondly, the principal demand that was imposed will teach us that $d\varphi$ depends linearly upon the shift $\overline{PP'}$ (with the components dx_i):

$$d\varphi = \sum_i \varphi_i dx_i .$$

There is no difference between the various points of the manifold in regard to the nature of the length determination that is based at each of them and its metric connection in its neighborhood.

The linear and quadratic fundamental forms:

$$d\varphi = \varphi_i dx_i \quad \text{and} \quad ds^2 = g_{ik} dx_i dx_k$$

describe the metric of the manifold relative to a reference system (= coordinate system + gauge). They remain invariant under coordinate transformation, while the second one will take on a factor α that is a positive, continuous function of position (viz., the “gauge ratio”) under a change of gauge, and the first one will be reduced by the total differential $d \ln \alpha$.

A metric manifold is also affinely connected with no further assumptions. That is true on the basis of the demand that a vector will remain congruent to the line segment that is determined by the vector under parallel displacement. That is the *fundamental fact of infinitesimal geometry*. If we clarify the process of the lowering of an index i for a system of numbers a^i (regardless of whether further indices do or do not appear in addition to i) once and for all by the equations:

$$a_i = g_{ij} a^j$$

(and the opposite process, by their inverses) then the affine connection of a metric manifold can be inferred from the formulas:

$$\Gamma_{i,kr} + \Gamma_{k,ir} = \frac{\partial g_{ik}}{\partial x_r} + g_{ik} \varphi_r \quad (\Gamma_{i,rs} = g_{ij} \Gamma_{rs}^j),$$

$$\Gamma_{r,ik} = \frac{1}{2} \left(\frac{\partial g_{ir}}{\partial x_k} + \frac{\partial g_{kr}}{\partial x_i} - \frac{\partial g_{ik}}{\partial x_r} \right) + \frac{1}{2} (g_{ir} \varphi_k + g_{kr} \varphi_i - g_{ik} \varphi_r).$$

Let us now recall a geometric notion: Two vectors \mathfrak{x} and η at P are called *orthogonal* to each other when the symmetric bilinear form $(\mathfrak{x} \cdot \eta)$ that belongs to the quadratic form \mathfrak{x}^2 vanishes for them; that reciprocal relationship is independent of the gauge factor.

Tensor calculus. – A (twice-covariant, simply-contravariant) *tensor* (of rank 3) at the point P is a linear form in the three series of variables ξ, η, ζ :

$$\sum_{i,k,l=0}^3 a^l_{ik} \xi^i \eta^k \zeta_l$$

that depends upon the coordinate system to which one refers the neighborhood of P , assuming that its dependency is of the following type: The expressions for the linear form in two coordinate system go to each other when one transforms the first two variables cogrediently and the last one contragrediently to the differentials [formula (2)]. The concept of a tensor is free of any relationship to the metric or affine connection of the manifold. Scalars fit into the system of tensors as tensors of rank 0. Tensors of rank 1 are called “vectors”; as before, we understand the term “vector” with no additional qualifier to mean a contravariant vector. The skew-symmetric covariant tensors play a special role and shall be called “linear tensors,” for brevity. The basic operations of *tensor algebra* by which only tensors at one and the same point P are coupled with each other are: addition, multiplication, and contraction; they do not assume that the manifold is either metrically or affinely connected. The same thing is still true for the analysis of *linear* tensors, which tells us how a linear tensor of rank $\nu + 1$ can be generally produced from one of rank ν by differentiation:

$$\frac{\partial u}{\partial x_i} = u_i, \quad \left| \frac{\partial u_i}{\partial x_k} - \frac{\partial u_k}{\partial x_i} = u_{ik} \right|, \quad \dots$$

However, the components of the affine connection will enter into the differentiation process of general *tensor analysis* (which is not restricted to the linear tensors). Hence, tensor analysis is first developed completely in an affinely-connected space. (By contrast, no metric is assumed). As an example, we mention:

$$\frac{\partial u^i}{\partial x_k} - \Gamma^i_{kr} u^r;$$

hence, a mixed tensor of rank 2 will arise from the vector field u^i .

If $\int \mathfrak{W} dx$ is an integral invariant (I shall write dx briefly for the integration element $dx_0 dx_1 dx_2 dx_3$) then \mathfrak{W} will be a function that depends upon the coordinate system that will get multiplied by absolute value of the functional determinant $|\alpha_i^k|$ under the transition from one coordinate system to the other. I shall refer to such a quantity as a “scalar density.” The concept of a *tensor density* (at the point P) is analogous: It is a linear form in several series of variables that depends upon the coordinate system, such that this linear form, as it reads in the coordinate system x_i , is converted into its expression in the coordinate system \bar{x}_i by multiply by the absolute value of the functional determinant and transforming the variables according to the same schema as above. The concept is free of any relationship to the metric or affine connection. The skew-symmetric contravariant tensor densities play a special role and shall be called linear tensor densities. *Tensors = intensities, tensor densities = quantities.* Whereas the distinction between these two types of quantities is blurred over in **Riemannian** geometry, here, we are in a position to make a sharp mathematical distinction between intensive and quantitative quantities. The basic operations of the *algebra of tensor densities* are: addition, multiplication of a tensor by a tensor density, and contraction; they assume neither a metric nor an affine connection. The same thing is true for the analysis of the *linear* tensor densities, which shows one how to produce a linear tensor density of rank $\nu - 1$ from one of rank ν by a process that has the character of a divergence:

$$\frac{\partial v^i}{\partial x_i} = v, \quad \left| \frac{\partial v^{ik}}{\partial x_i} = v^i \right|, \quad \dots$$

However, the components of the affine connection enter into the divergence and differentiation processes of the general *analysis of tensor densities*. For example:

$$\frac{\partial w_i^k}{\partial x_k} - \Gamma_{i\alpha}^\beta w_\beta^\alpha ;$$

in that way, a mixed tensor density w_i^k of rank 2 will arise from a vector density.

It is in the nature of tensors and tensor densities that the representative linear forms depend upon only the coordinate system but not the gauge, as well. However, we would also like to apply those names in an extended and figurative sense when the linear forms depend upon the coordinate systems in the way that was depicted above, but also depend upon the gauge, as well, and indeed, in such a way that they will be multiplied by a power a^e of the gauge ratio under re-gauging [viz., tensors (tensor densities, resp.) of *weight e*]. Just the same, we will regard this extension as only a tool that we would like to introduce for the sake of computational convenience. Namely, the following two operations will exist in the extended domain (of which one can naturally speak only in a metric manifold):

1. By lowering an index, the components of a tensor of weight e are converted into the components of a tensor of weight $e + 1$.

The character of that index will go from contravariant to covariant in that way. The converse is true for the raising of an index.

2. A tensor density of weight $e + 2$ will be produced by multiplying a tensor of weight e by \sqrt{g} ($-g$ is the determinant of the g_{ik} , so \sqrt{g} is the positive square root of that positive number g).

The latter operation shall be suggested once and for all by saying that one converts the Latin letters that denote a tensor into the corresponding German ones.

Curvature. – If a line segment advances congruently along a closed curve then upon its return to the starting point, it will not coincide with the starting line segment, in general. In order to find a measure of this “non-integrability” of the transfer of line segments, one performs a differential decomposition (precisely as one does with the line integral in **Stokes’s** theorem): One spans a surface with the closed curve, which one thinks of as being given by a parametric representation, and decomposes it into infinitely-small parallelograms by means of the coordinate lines. One must then determine the change ∇l that the length of a line segment experiences when the line segment goes around such a surface element that is spanned by the two elements dx_i and δx_i of the coordinate lines while remaining congruent to itself, and thus possesses the components:

$$\Delta x_{ik} = dx_i \delta x_k - dx_k \delta x_i .$$

One finds that:

$$\nabla l = -l \nabla \varphi ,$$

and the factor $\nabla \varphi$ in this will depend upon the surface element; namely, one has:

$$\nabla \varphi = f_{ik} dx_i \delta x_k = \frac{1}{2} f_{ik} \Delta x_{ik} , \quad f_{ik} = \frac{\partial \varphi_i}{\partial x_k} - \frac{\partial \varphi_k}{\partial x_i} .$$

We can correspondingly refer to the linear tensor f_{ik} of rank 2 that is determined uniquely by the metric as the *line segment curvature* of the metric manifold. Its vanishing is the necessary and sufficient condition for the transfer of length to be integrable, and therefore **Riemannian** geometry will be valid for the manifold.

The *vector curvature* has precisely the same relationship to the parallel displacement of vectors as the line segment curvature that was just constructed has to the congruent transplantation of line segments. The definition of vector curvature, which we shall also refer to casually as “curvature,” assumes only an affine connection on the manifold. An arbitrary vector \mathfrak{x} will suffer only a change $\nabla \mathfrak{x}$ while circumnavigating our infinitely-small surface element, which emerges from \mathfrak{x} by a linear map or “matrix” ∇F :

$$\nabla \mathfrak{x} = \nabla F(\mathfrak{x}), \quad \text{in components } \nabla \xi^\alpha = \nabla F_\beta^\alpha \cdot \xi^\beta .$$

∇F also depends linearly upon the surface element here:

$$\nabla F = F_{ik} dx_i \delta x_k = \frac{1}{2} F_{ik} \Delta x_{ik} \quad (F_{ki} = -F_{ik}).$$

For that reason, the curvature is best referred to as a “linear matrix tensor of rank 2.” However, if we look at the coefficients $F_{\beta ik}^\alpha$ of the matrix F_{ik} then the curvature will seem to be a tensor of rank four; one has:

$$(5) \quad F_{\beta ik}^\alpha = \left(\frac{\partial \Gamma_{\beta k}^\alpha}{\partial x_i} - \frac{\partial \Gamma_{\beta i}^\alpha}{\partial x_k} \right) + (\Gamma_{ri}^\alpha \Gamma_{\beta k}^r - \Gamma_{rk}^\alpha \Gamma_{\beta i}^r).$$

The vector curvature must include the line segment curvature as a component, since the parallel displacement of a vector automatically indeed carries with it the congruent transplantation of the line segment that it determines. In fact, if we decompose $\nabla \mathfrak{r}$ into a component $*\nabla \mathfrak{r}$ that is orthogonal to \mathfrak{r} and one that is parallel to \mathfrak{r} then we will get:

$$\nabla \mathfrak{r} = *\nabla \mathfrak{r} - \frac{1}{2} \mathfrak{r} \nabla \varphi.$$

Hand-in hand with that, there is a corresponding splitting of the curvature:

$$(6) \quad F_{\beta ik}^\alpha = *F_{\beta ik}^\alpha - \frac{1}{2} \delta_\beta^\alpha f_{ik},$$

whose first component must consequently be called the “direction curvature.” The numbers $*F_{\beta ik}^\alpha$ are not only skew-symmetric in the indices i and k , but also in the α and β .

For later calculations, we shall use the tensor $F_{i\alpha k}^\alpha = F_{ik}$ that arises by contraction and the scalar of weight -1 , $F_i^i = F$, that arises from it by another contraction. The **Riemannian** curvature quantities that emerge from them by setting the φ_i to zero might be denoted by $-R_{ik}$ ($-R$, resp.). One then has:

$$(7) \quad -F = R + \frac{3}{\sqrt{g}} \frac{\partial(\sqrt{g} \varphi^i)}{\partial x_i} + \frac{3}{2} (\varphi_i \varphi^i).$$

The linear tensor density \mathfrak{f}^{ik} (of weight 0) will arise from the linear tensor f_{ik} (in the *four-dimensional* world), and one will get the scalar density:

$$\mathfrak{l} = \frac{1}{4} f_{ik} \mathfrak{f}^{ik}$$

from both of them. $\int \mathfrak{l} dx$ is the simplest integral invariant that one can construct from the metric, and an integral invariant of such a simple structure can exist only in a four-dimensional manifold. Naturally, the integral $\int \sqrt{g} dx$ that appears in **Riemannian** geometry has no meaning here.

The static case. – The metric field in the four-dimensional world is a static one when one can choose the coordinate system and gauge in such a way that the fundamental linear form is equal to φdx_0 , and the quadratic form is equal to $c^2 dx_0^2 - d\sigma^2$. In them, φ and $c (> 0)$ are functions of only x_1, x_2, x_3 , and $d\sigma^2$ is a positive-definite quadratic form in the variables x_1, x_2, x_3 . x_0 is the *time coordinate*, and x_1, x_2, x_3 are the *space coordinates*. The special form of the fundamental form will not be perturbed by coordinate transformations and changes of gauge only when the time coordinate x_0 suffers a linear transformation in its own right, and, at the same time, the gauge ratio is constant. In the static case, we then get a three-dimensional *Riemannian space* with the fundamental metric form $d\sigma^2$ and the two scalar fields c and φ in that space. One then chooses the units of length and time (cm, sec) as arbitrary units of measurement. $d\sigma^2$ has the dimension cm^2 , the speed of light c has the dimension $\text{cm} \cdot \text{sec}^{-1}$, and φ has the dimension sec^{-1} . Namely, one must observe that the three-dimensional space is presented, not as an arbitrary metric space (in which the transfer of line elements proves to be non-integrable), but as a **Riemannian** one.

Chapter II. Field laws and conservation laws.

Transition to physics. – The special theory of relativity teaches that the world-geometry that prevails in the four-dimensional world is not based upon a “Galilean” metric, but a “Euclidian one.” However, a disharmony arises from this, since the *laws of local action* of modern physics will then have **Euclidian distance geometry** at their foundations. In this, one can glimpse a speculative basis for replacing the **Euclidian** world-geometry with the **Riemannian** one and ultimately with the pure local geometry that was just discussed. **Einstein** remains rooted in **Riemannian** geometry. However, two more notions are characteristic of his “general theory of relativity,” along with the transition from **Euclidian** distance geometry to **Riemannian** local geometry:

1. The metric is not given *a priori*, but depends upon the distribution of matter.

In connection with this, the *relativity of motion* is the only argument by which the theory becomes persuasive.

2. The properties of gravitation (e.g., equality of gravitational and inertial mass) that are known from experiment and not understood up to now will become tangible when one attributes the gravitational phenomena to the deviation of the metric from the **Euclidian** one, and not to certain forces that act “in” in the metric world.

Although on first glance its structure deviates completely from the **Newtonian** theory, the theory of gravitation that comes about in that way is in complete agreement with all astronomical experiments, as one will see by pursuing its consequences under certain simplifying assumptions.

The new extension that is presented here is likewise initially concerned with only the world-geometric foundations of physics, and as such, represents a consistent expansion of relativistic ideas. However, with just the same power that the relativity of motion

compels us towards **Einstein**'s theory, the belief in the *relativity of magnitude* will compel us to take that additional step. Just as we were given gravitation by the former theory, we are now given *electromagnetism*. If we combine the potentials of the gravitational field into a quadratic differential form, as **Einstein** did, then we will know that the potentials of the electromagnetic field define the coefficients of an invariant linear differential form. Therefore, it stands to reason that one can identify the fundamental linear form that appears along with the quadratic one in the pure local geometry with the potential form of the electromagnetic field. Not only the gravitational forces, but also the electromagnetic ones, would then arise from the world metric, and since other truly primordial force effects than those two are simply not known to us, the theory that would emerge in that way would fulfill the dream of **Descartes** of a purely-geometrical physics in a remarkable way that was, admittedly, not at all foreseen by him by showing: The conceptual content of physics does not overlap with geometry in any way, *but only the metric field manifests itself in matter and natural forces*. Gravitation and electricity would then be explained by a unifying source. The entire wealth of experiments by which **Maxwell**'s theory is established speaks for those ideas. Here (in infinitesimal geometry), as there (in **Maxwell**'s theory), the linear form $\varphi_i dx_i$ is then determined only up to an additive total differential, and it is only the "field" (= line segment curvature) that is derived from it:

$$f_{ik} = \frac{\partial \varphi_i}{\partial x_k} - \frac{\partial \varphi_k}{\partial x_i},$$

and which satisfies the equations:

$$\frac{\partial f_{kl}}{\partial x_i} + \frac{\partial f_{li}}{\partial x_k} + \frac{\partial f_{ik}}{\partial x_l} = 0,$$

that will be free of arbitrariness. The electromagnetic quantity of action that **Maxwell**'s theory obeys:

$$\int \! \! \! \int dx = \frac{1}{4} \int f_{ik} f^{ik} dx,$$

is also obtained here as an invariant, and in fact, as the simplest integral invariant that exists at all. Not only does it lead to a deeper understanding of **Maxwell**'s theory, but even the fact that the world is four-dimensional, which was always accepted as "coincidental" up to now, will become understandable. The cited basis seems to me to be perhaps rigorously equivalent to the one that led **Einstein** to with his general theory of relativity, and its speculative character might also emerge as even more blatant to us.

At first, it might be suspicious ⁽¹⁾ that in pure local geometry the transfer of line segments is not integrable when an electromagnetic field is present. Is that not in flagrant contradiction to the behavior of rigid bodies and clocks? However, the functioning of those measuring instruments is a physical process whose evolution is determined by the laws of nature, and as such has nothing to do with the ideal process of the "congruent transplanting of world-line segments" that we appealed to as the mathematical structure

⁽¹⁾ For an objection to the theory that is proposed here that Einstein formulated, cf., the Addendum to the author's aforementioned note to the Akademie.

of world-geometry. Even in the special theory of relativity, the connection between the metric field and the behavior of yardsticks and clocks will become entirely opaque as soon as one does not restrict oneself to quasi-stationary motion. Hence, those instruments also play a role that is indispensable in practice as indicators of the metric field (simpler processes – e.g., the spreading of light – would be theoretically preferable for that purpose), so it would obviously be wrong to *define* the metric field by the information that is extracted from them directly. We must return to that question after we have presented the laws of nature.

The implementation of the theory must show whether it is confirmed. The **Maxwell-Lorentz** theory was characterized by the duality of matter and the electromagnetic field. That was eliminated by **Mie**'s theory ⁽¹⁾ (based upon the special theory of relativity). However, the juxtaposition of the electromagnetic field (“matter in the extended sense,” as **Einstein** said) with the gravitational field appeared in its place by considering gravitation. He showed that most clearly by splitting the **Hamilton** function that **Einstein**'s theory is based upon into two pieces ⁽²⁾. That splitting will also be avoided in our theory. The integrand of the quantity of action $\int \mathfrak{W} dx$ must be a scalar density that arises from the metric, and the laws of nature must be summarized in **Hamilton**'s principle: For any infinitesimal change δ of the world-metric that vanishes outside of a finite domain, the change in the total quantity of action:

$$\delta \int \mathfrak{W} dx = \int \delta \mathfrak{W} dx$$

will be equal to zero. (The integral extends over all of the world, or – what amounts to the same thing – over a finite domain, outside of which the variation δ vanishes.) The quantity of action is necessarily a pure number in our theory; indeed, it could not be otherwise if a quantum of action were to exist. We assume that \mathfrak{W} is a second-order expression; i.e., it is, on the one hand, constructed from the g_{ik} and their derivatives of first and second order, and on the other hand, from the φ_i and their first-order derivatives. The simplest example is the **Maxwell** action density \mathfrak{L} . However, in this chapter, we would not like to make any special Ansatz for \mathfrak{W} , but to investigate what one can conclude from only the fact that $\int \mathfrak{W} dx$ is a coordinate-invariant and gauge-invariant integral. We shall then appeal to a method that was given by **F. Klein** ⁽³⁾.

Consequences of the invariance of the quantity of action.

a) *Gauge invariance.* If we give infinitely-small increases $\delta\varphi_i$, δg_{ik} to the quantities φ_i , g_{ik} that describe the metric relative to a reference system, and if \mathfrak{X} means a finite world-domain then the effect of partial integration is that the integral of the associated

⁽¹⁾ Ann. Phys. (Leipzig) **37**, **39**, **40** (1912/13).

⁽²⁾ Cf., **Einstein**, “Hamilton's Prinzip und allgemeine Relativitätstheorie,” Sitz. Preuss. Akad. Wiss. (1916), pp. 1111.

⁽³⁾ Nachr. d. Ges. d. Wiss. zu Göttingen, Session on 19 July 1918.

change $\delta\mathfrak{W}$ in \mathfrak{W} over the domain \mathfrak{X} will split into two parts: a divergence integral and an integral whose integrand is only a linear combination of $\delta\varphi_i$ and δg_{ik} :

$$(8) \quad \int_{\mathfrak{X}} \delta\mathfrak{W} dx = \int_{\mathfrak{X}} \frac{\partial(\delta v^k)}{\partial x_k} dx + \int_{\mathfrak{X}} (\mathfrak{w}^i \delta\varphi_i + \frac{1}{2} \mathfrak{W}^{ik} \delta g_{ik}) dx \quad \{\mathfrak{W}^{ki} = \mathfrak{W}^{ik}\}.$$

In this, \mathfrak{w}^i , δv^i are the components of contravariant vector densities, but \mathfrak{W}_k^i is a mixed tensor density of rank 2 (in the proper sense). The components δv^i are linear combinations of:

$$\delta\varphi_i, \delta g_{ik} \quad \text{and} \quad \delta g_{ik,r} \quad \left\{ g_{ik,r} = \frac{\partial g_{ik}}{\partial x_r} \right\}.$$

We now express the idea that $\int_{\mathfrak{X}} \mathfrak{W} dx$ does not change when the gauge of the world changes infinitesimally. If $\alpha = 1 + \pi$ is the ratio of the varied gauge to the original one then π will be an infinitesimal scalar field that characterizes the process that can be given arbitrarily. The fundamental quantities will experience the increments:

$$(9) \quad \delta g_{ik} = \pi \cdot g_{ik}, \quad \delta\varphi_i = -\frac{\partial\pi}{\partial x_i}$$

under that process. If we substitute these values in δv^i then the expression:

$$(10) \quad \mathfrak{s}^k(\pi) = \pi \cdot \mathfrak{s}^k + \frac{\partial\pi}{\partial x_\alpha} \cdot \mathfrak{h}^{k\alpha}$$

might emerge. The variation (8) of the action integral must vanish for (9): We then formulate the fact of gauge invariance as:

$$\int_{\mathfrak{X}} \frac{\partial \mathfrak{s}^k(\pi)}{\partial x_k} dx + \int_{\mathfrak{X}} \left(-\mathfrak{w}^i \frac{\partial\pi}{\partial x_i} + \frac{1}{2} \mathfrak{W}_i^i \pi \right) dx = 0.$$

If one converts the first term in the second integral by partial integration then one can write:

$$(11) \quad \int_{\mathfrak{X}} \frac{\partial(\mathfrak{s}^k(\pi) - \pi \mathfrak{w}^k)}{\partial x_k} dx + \int_{\mathfrak{X}} \pi \left(\frac{\partial \mathfrak{w}^i}{\partial x_i} + \frac{1}{2} \mathfrak{W}_i^i \right) dx = 0$$

in place of this. That will next imply the identity:

$$(12) \quad \frac{\partial \mathfrak{w}^i}{\partial x_i} + \frac{1}{2} \mathfrak{W}_i^i = 0$$

in the way that is well-known from the calculus of variations: If this function of position were non-zero at a location (x_i) – say, positive – then one could delimit a neighborhood \mathfrak{X} of that location that is small enough that the function would remain positive in all of \mathfrak{X} . If one chooses that domain to be \mathfrak{X} in (11), but π is a function that vanishes outside of \mathfrak{X} and is consistently ≥ 0 inside of \mathfrak{X} then the first integral will vanish, but the second one will prove to be positive, which contradicts equation (11). Once that is known, (11) will yield the further equation:

$$\int_{\mathfrak{X}} \frac{\partial(\mathfrak{s}^k(\pi) - \pi \mathfrak{w}^k)}{\partial x_k} dx = 0.$$

For a given scalar field π , this must be true for any finite domain \mathfrak{X} , and as a result, one must have:

$$(13) \quad \frac{\partial(\mathfrak{s}^k(\pi) - \pi \mathfrak{w}^k)}{\partial x_k} = 0.$$

If we substitute this in (10) and observe that the values of:

$$\pi, \quad \frac{\partial \pi}{\partial x_i}, \quad \frac{\partial^2 \pi}{\partial x_i \partial x_k}$$

can be given arbitrarily at a location then that formula will split into the following identities:

$$1. \quad \frac{\partial \mathfrak{s}^k}{\partial x_k} = \frac{\partial \mathfrak{w}^k}{\partial x_k}, \quad 2. \quad \mathfrak{s}^i + \frac{\partial \mathfrak{h}^{\alpha i}}{\partial x_\alpha} = \mathfrak{w}^i, \quad 3. \quad \mathfrak{h}^{\alpha\beta} + \mathfrak{h}^{\beta\alpha} = 0.$$

Since $\partial \pi / \partial x_i$ are the components of a covariant vector field that arises from the scalar field π , the fact that $\mathfrak{s}^i(\pi)$ is a vector density will imply that \mathfrak{s}^i is a vector density, and \mathfrak{h}^{ik} is a tensor density, and indeed, from 3, a linear tensor density of rank 2. Due to the skew-symmetry of \mathfrak{h} , (1.) will be a consequence of (2.), since one has:

$$\frac{\partial^2 \mathfrak{h}^{\alpha\beta}}{\partial x_\alpha \partial x_\beta} = 0.$$

b) Coordinate invariance. We subject the world-continuum to an infinitesimal deformation, under which the individual point (x_i) will experience a shift with the components $\xi^i(x)$; the metric will be unchanged under the deformation. δ will denote the change in any quantity that is affected by the deformation when one remains at the

same space-time location, while δ' will be its change when one displaces the space-time location along with it. One has:

$$(14) \quad \left\{ \begin{array}{l} -\delta\varphi_i = \left(\varphi_r \frac{\partial \xi^r}{\partial x_i} + \frac{\partial \varphi_r}{\partial x_i} \xi^r \right) + \frac{\partial \pi}{\partial x_i}, \\ -\delta g_{ik} = \left(g_{ir} \frac{\partial \xi^r}{\partial x_k} + g_{kr} \frac{\partial \xi^r}{\partial x_i} + \frac{\partial g_{ik}}{\partial x_r} \xi^r \right) - \pi g_{ik}. \end{array} \right.$$

In this, π means an infinitesimal scalar field that is unaffected by the conditions that we have established. The invariance of the quantity of action under coordinate transformations and changes of gauge is expressed by this variational formula (which includes five arbitrary functions ξ^i and π):

$$(15) \quad \delta' \int_{\mathfrak{x}} \mathfrak{W} dx = \int_{\mathfrak{x}} \left\{ \frac{\partial(\mathfrak{W} \xi^i)}{\partial x_k} + \delta \mathfrak{W} \right\} dx = 0.$$

If one wishes to express only the coordinate invariance of that expression then one must choose $\pi = 0$; however, the variational formulas (14) would have no invariant character. In fact, that condition would mean: The two fundamental forms shall be varied by the deformation in such a way that length l of a line element that is carried by the deformation must remain invariant: $\delta' l = 0$. However, that equation does not express the process of the congruent transplanting of a line segment, as opposed to:

$$\delta' l = -l (\varphi_i \delta' x_i) = -l (\varphi_i \xi^i).$$

If invariant formulas are to come about then we must not choose $\pi = 0$ in (14), but $\pi = -(\varphi_i \xi^i)$, namely:

$$(16) \quad \left\{ \begin{array}{l} -\delta\varphi_i = f_{ir} \xi^r, \\ -\delta g_{ik} = \left(g_{ir} \frac{\partial \xi^r}{\partial x_k} + g_{kr} \frac{\partial \xi^r}{\partial x_i} \right) + \left(\frac{\partial g_{ik}}{\partial x_r} + g_{ik} \varphi_r \right) \xi^r. \end{array} \right.$$

The variation of the two fundamental forms that it expresses is such that *the metric is carried along unchanged by the deformation and every line element will be transplanted congruently*. One can also recognize the invariant character of equations (16) analytically. It will be revealed in the second one when one introduces the mixed tensor:

$$\frac{\partial \xi^i}{\partial x_k} - \Gamma_{kr}^i \xi^r = \xi_k^i.$$

It will then read:

$$-\delta g_{ik} = \xi_{ik} + \xi_{ki}.$$

Once the gauge invariance is exploited under a), it will suffice to make any particular choice for π ; from the standpoint of invariance, the $\pi = -(\varphi_i \xi^i)$ that leads to (16) is the only one possible.

For the variation (16), let:

$$\mathfrak{W}\xi^i + \delta v^k = \mathfrak{G}^k(\xi).$$

$\mathfrak{G}^k(\xi)$ is a linear differential vector density that depends upon the arbitrary vector field ξ^i ; I write it explicitly:

$$\mathfrak{G}^k(\xi) = \mathfrak{G}_i^k \frac{\partial \xi^i}{\partial x_k} + \bar{\mathfrak{H}}_i^{k\alpha} \frac{\partial \xi^i}{\partial x_\alpha} + \frac{1}{2} \mathfrak{H}_i^{k\alpha\beta} \frac{\partial^2 \xi^i}{\partial x_\alpha \partial x_\beta}.$$

(Naturally, the last coefficient is symmetric in the indices $\alpha\beta$.) If we introduce the expressions (8), (16) into (15) then that will produce an integral whose integrand reads:

$$\frac{\partial \mathfrak{G}^k(\xi)}{\partial x_k} - \mathfrak{W}_i^k \frac{\partial \xi^i}{\partial x_k} - \left\{ f_{ki} \mathfrak{w}^k + \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial x_i} + g_{\alpha\beta} \varphi_i \right) \right\} \mathfrak{W}^{\alpha\beta}.$$

Since:

$$\frac{\partial g_{\alpha\beta}}{\partial x_i} + g_{\alpha\beta} \varphi_i = \Gamma_{\alpha, \beta i} + \Gamma_{\beta, \alpha i},$$

and due to the symmetry of $\mathfrak{W}^{\alpha\beta}$, one will have:

$$\frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial x_i} + g_{\alpha\beta} \varphi_i \right) \mathfrak{W}^{\alpha\beta} = \Gamma_{\alpha, \beta i} \mathfrak{W}^{\alpha\beta} = \Gamma_{\beta i}^\alpha \mathfrak{W}_\alpha^\beta.$$

If we perform yet another partial integration on the second term of our integrand then we will get:

$$\int_{\bar{x}} \frac{\partial(\mathfrak{G}^k(\xi) - \mathfrak{W}_i^k \xi^i)}{\partial x_k} dx + \int_{\bar{x}} \left(\frac{\partial \mathfrak{W}_i^k}{\partial x_k} - \Gamma_{\beta i}^\alpha \mathfrak{W}_\alpha^\beta + f_{ik} \mathfrak{w}^k \right) \xi^i dx = 0.$$

The identities:

$$(17) \quad \left(\frac{\partial \mathfrak{W}_i^k}{\partial x_k} - \Gamma_{\beta i}^\alpha \mathfrak{W}_\alpha^\beta \right) + f_{ik} \mathfrak{w}^k = 0$$

and

$$(18) \quad \frac{\partial(\mathfrak{G}^k(\xi) - \mathfrak{W}_i^k \xi^i)}{\partial x_k} = 0$$

arise from this, by the argument that was applied above. The last one decomposes into the four following ones:

$$\begin{aligned} \text{I.} \quad & \frac{\partial \mathfrak{S}_i^k}{\partial x_k} = \frac{\partial \mathfrak{W}_i^k}{\partial x_k}, & \text{II.} \quad & \mathfrak{S}_i^k + \frac{\partial \bar{\mathfrak{H}}_i^{\alpha k}}{\partial x_\alpha} = \mathfrak{W}_i^k, \\ \text{III.} \quad & (\bar{\mathfrak{H}}_i^{\alpha\beta} + \bar{\mathfrak{H}}_i^{\beta\alpha}) + \frac{\partial \mathfrak{H}_i^{\gamma\alpha\beta}}{\partial x_\gamma} = 0, & \text{IV.} \quad & \mathfrak{H}_i^{\alpha\beta\gamma} + \mathfrak{H}_i^{\beta\gamma\alpha} + \mathfrak{H}_i^{\gamma\alpha\beta} = 0. \end{aligned}$$

If one replaces:

$$\mathfrak{H}_i^{\gamma\alpha\beta} \text{ with } -\mathfrak{H}_i^{\alpha\beta\gamma} - \mathfrak{H}_i^{\beta\alpha\gamma}$$

in III and IV then it will emerge that:

$$\bar{\mathfrak{H}}_i^{\alpha\beta} - \frac{\partial \mathfrak{H}_i^{\alpha\beta\gamma}}{\partial x_\gamma} = \mathfrak{H}_i^{\alpha\beta}$$

is skew-symmetric in the indices $\alpha\beta$. If we introduce $\mathfrak{H}_i^{\alpha\beta}$, in place of $\bar{\mathfrak{H}}_i^{\alpha\beta}$, then III and IV will include merely statements of symmetry, but II will go to:

$$\text{(II)*} \quad \mathfrak{S}_i^k + \frac{\partial \mathfrak{H}_i^{\alpha k}}{\partial x_\alpha} + \frac{\partial^2 \mathfrak{H}_i^{\alpha\beta k}}{\partial x_\alpha \partial x_\beta} = \mathfrak{W}_i^k.$$

It follows from this, because due to the symmetry conditions:

$$\frac{\partial^2 \mathfrak{H}_i^{\alpha\beta k}}{\partial x_\alpha \partial x_\beta} = 0, \quad \text{so one will have} \quad \frac{\partial^3 \mathfrak{H}_i^{\alpha\beta\gamma}}{\partial x_\alpha \partial x_\beta \partial x_\gamma} = 0.$$

The invariance character of the coefficients \mathfrak{S} and \mathfrak{H} of $\mathfrak{S}^k(\xi)$ (in particular, that of the quantities \mathfrak{S}_i^k) can be described most simply and completely by the saying that $\mathfrak{S}^k(\xi)$ is a vector density (but ξ^i is a vector). It will then emerge from this that \mathfrak{S}_i^k are *not* the components of a mixed tensor density; in this case, we speak of a “pseudo-tensor density.”

Example. For $\mathfrak{W} = \mathfrak{l}$, as one sees immediately, one has:

$$\delta v^k = \mathfrak{f}^{ik} \delta \varphi_i,$$

so as a result:

$$\mathfrak{s}^i = 0, \quad \mathfrak{h}^{ik} = \mathfrak{f}^{ik}, \quad \mathfrak{S}_i^k = \delta_i^k \mathfrak{l} - f_{i\alpha} \mathfrak{f}^{ik}, \quad \text{and the quantities } \mathfrak{H} = 0.$$

Our identities then imply that:

$$\mathfrak{w}^i = \frac{\partial f^{\alpha i}}{\partial x_\alpha}, \quad \frac{\partial \mathfrak{w}^i}{\partial x_i} = 0, \quad \mathfrak{W}_i{}^i = 0,$$

$$\mathfrak{W}_i{}^k = \mathfrak{S}_i{}^k, \quad \left(\frac{\partial \mathfrak{S}_i{}^k}{\partial x_k} - \frac{1}{2} \frac{\partial g_{\alpha\beta}}{\partial x_i} \mathfrak{S}^{\alpha\beta} \right) + f_{i\alpha} \frac{\partial f^{\beta\alpha}}{\partial x_\beta} = 0.$$

The two formulas in the last row are confirmed by calculation in **Maxwell's** theory. In it, the components $\mathfrak{S}_i{}^k$ define the tensor density of the energy of the electromagnetic field, and the last equation says that the ponderomotive force will arise from that tensor density by taking its divergence.

Field laws and conservation laws. – If one takes δ in (8) to be an arbitrary variation that vanishes outside of a finite domain and takes \mathfrak{X} to be the entire world or a domain such that $\delta = 0$ outside of it then that will give:

$$\int \delta \mathfrak{W} dx = \int (\mathfrak{w}^i \delta \varphi_i + \frac{1}{2} \mathfrak{W}^{ik} \delta g_{ik}) dx.$$

It emerges from this that the following invariance laws are included in Hamilton's principle $\int \delta \mathfrak{W} dx = 0$:

$$\mathfrak{w}^i = 0, \quad \mathfrak{W}_i{}^k = 0.$$

The first of these are the *electromagnetic laws*, while the second ones are the *gravitational laws*. Five identities exist between the left-hand sides of these equations that are specified above in (12) and (17). Five of the equations in the system of field equations are therefore superfluous, corresponding to the transition from a reference system to any other one, which depends upon five arbitrary functions. s^i is the vector density of electric four-current, $\mathfrak{S}_i{}^k$ is the pseudo-tensor of energy, and \mathfrak{h}^{ik} is the electromagnetic field density. In the case of **Maxwell's** theory, which is indeed only valid in the ether (as it must be), $s^i = 0$, $\mathfrak{h}^{ik} = \mathfrak{f}^{ik}$, and the $\mathfrak{S}_i{}^k$ are the classical expressions. From (1.) and (I.), one generally has the conservation laws:

$$\frac{\partial s^i}{\partial x_i} = 0, \quad \frac{\partial \mathfrak{S}_i{}^k}{\partial x_k} = 0.$$

Indeed, the conservation laws follow from the field laws in two ways; namely, not only does:

$$\frac{\partial s^i}{\partial x_i} \equiv \frac{\partial \mathfrak{w}^i}{\partial x_i}, \quad \text{but it also} \equiv -\frac{1}{2} \mathfrak{W}_i{}^i,$$

and not only does:

$$\frac{\partial \mathfrak{S}_i^k}{\partial x_k} \equiv \frac{\partial \mathfrak{W}_i^k}{\partial x_k}, \quad \text{but it also} \equiv \Gamma_{i\beta}^\alpha \mathfrak{W}_i^{\beta} - f_{ik} \mathfrak{w}^k.$$

The close relationship that exists between the conservation laws for energy-impulse and the coordinate invariance has been pursued by various authors in **Einstein's** theory (¹). However, a fifth conservation law gets added to these four, namely, the conservation of electricity, and it must consequently correspond to an invariance property that also brings with it a fifth arbitrary function; in our theory, it is seen to be gauge invariance. Moreover, the older investigations into the energy-impulse theorem never led to an entirely transparent result. One then makes no special assumption about the quantity of action in **Einstein's** theory that could, however, appeal in any way to the law of conservation of energy and impulse, since it does not reduce to it in the classical cases. That has left me very uneasy for quite some time already. However, we get its complete explanation here: One must couple the coordinate invariance with the gauge invariance in such a way that our theory – viz., formula (16) – will be implied by that in its own right in order to lead to the correct conservation laws. That complete connection is obviously a very strong argument for the validity of our theory, namely, that the laws of nature are not only coordinate-invariant, but also gauge-invariant.

Let us add this: From (2.) [viz., the equations into which (13) decompose], the electromagnetic equations read as follows:

$$\frac{\partial \mathfrak{h}^{ik}}{\partial x_k} = \mathfrak{s}^i \quad \left(\text{and } \frac{\partial f_{kl}}{\partial x_i} + \frac{\partial f_{li}}{\partial x_k} + \frac{\partial f_{ik}}{\partial x_l} = 0 \right).$$

Without specializing the quantity of action, we can read off the entire structure of Maxwell's theory from gauge invariance alone. The only laws that will be affected by the special form of the **Hamilton** function \mathfrak{W} are the ones by which the current \mathfrak{s}^i and field density \mathfrak{h}^{ik} are determined from the fundamental quantities φ_i, g_{ik} .

From (13) and (18), the field laws and the conservation laws that belong to them can be summarized most clearly in the two simple equations:

$$\frac{\partial \mathfrak{s}^i(\pi)}{\partial x_i} = 0, \quad \frac{\partial \mathfrak{S}^i(\xi)}{\partial x_i} = 0$$

(viz., the **Hilbert-Klein** form of the field laws).

(¹) Such as **H. A. Lorentz, Hilbert, Einstein, Klein**, and the author.

Chapter III – Implementing a special action principle.

The Ansatz for \mathfrak{W} . – I shall base the further discussion upon the action principle that allows one to survey its analytical consequences most easily:

$$\mathfrak{W} = -\frac{1}{4}F^2\sqrt{g} + \beta \mathfrak{l}.$$

The meanings of \mathfrak{l} and F are gathered from the foregoing, while the constant β is a pure number. That gives:

$$\delta\mathfrak{W} = -\frac{1}{2}F\delta(F\sqrt{g}) + \frac{1}{4}F^2\delta\sqrt{g} + \beta\delta\mathfrak{l}.$$

It will simplify the calculations greatly when we fix the gauge of the world uniquely by the requirement that $-F$ must be equal to a (given positive) constant α ; that is possible, because F is an invariant of weight -1 . In that way, we arrive at the fact that the field laws are second-order differential equations. If we drop a divergence:

$$\delta\frac{\partial(\sqrt{g}\varphi^i)}{\partial x_i},$$

which will indeed vanish upon integration over the world, then $\delta\mathfrak{W}$ will become:

$$\delta\left(\beta\mathfrak{l} + \frac{\alpha^2\sqrt{g}}{4} - \frac{3\alpha\sqrt{g}}{4}(\varphi_i\varphi^i) - \frac{\alpha\sqrt{g}}{2}R\right).$$

If we then divide by α , set $\beta/\alpha = \lambda$, and convert the world-integral of $\delta(\frac{1}{2}R\sqrt{g})$ into the integral of $\delta\mathfrak{G}$ by partial integration, in which \mathfrak{G} depends upon only the g_{ik} and their first derivatives (¹), then we will come to the action principle:

$$(19) \quad \delta\int\left\{\lambda\mathfrak{l} - \mathfrak{G} + \frac{\alpha - 3(\varphi_i\varphi^i)}{4}\sqrt{g}\right\}dx = 0.$$

The structure of the integrand is clear: $\lambda\mathfrak{l}$ and $-\mathfrak{G}$ are the classical terms in **Maxwell's** theory of electromagnetism and **Einstein's** theory of gravitation. The “cosmological term” $(\alpha/4)\sqrt{g}$ is included in it, which appears here of necessity, along with the simplest term that can be added to the **Maxwellian** action density according to **Mie's** theory and which the existence of matter makes possible: i.e., $(\varphi_i\varphi^i)\sqrt{g}$. In this, one must observe that according to our theory, this Ansatz is one of a very limited number of possibilities (cf., the conclusion of the paper on this), and is, in any case, the only one that

(¹) \mathfrak{G} is the quantity that was denoted by $\frac{1}{2}\mathfrak{G}^*$ on pp. 114 of the cited paper of **Einstein**.

leads to differential equations of order no higher than two. In particular, it is not at all up to our discretion to assign the sign of the term $(\varphi_i \varphi^i)$ to be anything but the one in (19). From what was said, it is already clear that the principle (19) agrees with the laws of the electromagnetic and gravitational field that are accessible to experimental confirmation outside of matter.

Varying the φ_i will yield **Maxwell's** equations:

$$(20) \quad \frac{\partial f^{ik}}{\partial x_k} = -\frac{3}{2\lambda} \sqrt{g} \varphi^i.$$

The electromagnetic field density is therefore equal to f^{ik} here, and the expression on the right-hand side is the current density s^i . The divergence equation:

$$(21) \quad \frac{\partial(\sqrt{g} \varphi^i)}{\partial x_i} = 0$$

follows from this. Varying the g_{ik} will yield the gravitational equation:

$$(22) \quad -R_{ik} + \rho g_{ik} = \frac{3}{2} \varphi_i \varphi_k + \lambda S_{ik}^*,$$

in which S_{ik}^* are the components of the **Maxwellian** energy-impulse, and:

$$\rho = \frac{1}{2} R + \frac{-\alpha + 3(\varphi_i \varphi^i)}{4}.$$

If we contract then it will follow that:

$$R - \alpha + \frac{3}{2}(\varphi_i \varphi^i) = 0 \quad \text{and from that} \quad \rho = \frac{\alpha}{4}.$$

Since $-F = \alpha$, the first relation will once more yield (21) – viz., the conservation of electricity – which, as one confirms, is a double consequence of the field laws. The right-hand side of (22) is equal to:

$$\lambda (S_{ik}^* - \varphi_i s_k),$$

in complete agreement with **Mie's** theory. In the ether, the first term outweighs the second one, which is relevant only in the interior of material particles (e.g., atomic nuclei or electrons).

Our theory is based upon a certain unit of electricity. I call:

$$\frac{e\sqrt{\kappa}}{c_0}$$

(κ is **Einstein's** gravitational constant, and c_0 is the speed of light in ether) the *gravitational radius* of the charge e , so one can characterize this unit, as would follow from (22) by saying: It is the charge whose gravitational radius is equal to $\sqrt{\frac{1}{2}\lambda}$. That length must certainly be enormous, since otherwise equation (20) would contradict experiments; when the number $\beta = 1$, it will have the order of magnitude of the radius of the world. Our unit of electricity, and likewise the unit of action, will then be of cosmic magnitude, in any event. *The "cosmological" moment that Einstein first introduced into his theory heuristically is attached to our theory as a result of its first principles.*

Let us make two remarks about the static case! The static world is gauged inherently. (cf., Chap. I) One asks whether its natural gauge $F = \text{const.}$ is valid as a result of that. The answer is yes. If we gauge the world by the demand that $F = \text{const.}$ then the fundamental metric form will take on the factor F , and $d\varphi = \varphi dx_0$ must be replaced with:

$$\varphi dx_0 - \frac{dF}{F}.$$

Equation (21) will then imply that:

$$\frac{\partial \mathfrak{F}^1}{\partial x_1} + \frac{\partial \mathfrak{F}^2}{\partial x_2} + \frac{\partial \mathfrak{F}^3}{\partial x_3} = 0 \quad \left(F_i = \frac{\partial F}{\partial x_i} \right),$$

and it will follow from this that $F = \text{const.}$

The second remark is this: In the static case, the (00)th gravitational equation (22) reads:

$$c \left(\Delta c + \frac{\alpha}{4} c \right) = \frac{3}{2} \varphi^2 + \lambda S_{00}^*.$$

In this, Δ is the spatial **Poisson** operator that belongs to the fundamental metric form $d\sigma^2$. The right-hand side is positive here. Our action principle actually leads to a positive mass, and therefore an attractive force between them, a not repulsive one.

Mechanics – The Ansätze that originate in the presence of substances, and by which one accomplishes the transition from the energy-impulse principle to the mechanical equations that govern the motion of a material particle, prove to be impossible in our theory, since they contradict the required invariance properties. Moreover, as I have remarked here in passing, they will lead to a false value of the mass, just as they do in **Einstein's** theory on the very same basis and for the sake of which we must dismiss them completely here. The only tenable path that can lead to an actual derivation of the mechanical equations when one assumes the existence of material particles was proposed in Part 3 of **Mie's** trailblazing "Grundlagen einer Theorie der Materie" ⁽¹⁾, and was recently trodden by **Einstein** in order to prove the integral conservation laws for an isolated system ⁽²⁾. One imagines the material particles as being in a bounded volume Ω

⁽¹⁾ Ann. Phys. (Leipzig) **40** (1913), pp. 1.

⁽²⁾ Sitz. Preuss. Akad. Wiss. (1918).

whose dimensions are large compared to the actual nucleus of concentration of the particles, but small in comparison to any dimensions in which the external field varies noticeably. By its motion, Ω will describe a channel in whose interior the current filament flows. The coordinate system, which consists of the “time coordinate” $x_0 = t$ and the “spatial coordinates” x_1, x_2, x_3 are so arranged that the “space” $x_0 = \text{const.}$ cuts through the channel (the cross-section is the aforementioned volume). The pseudo-tensor density of total energy will be denoted by \mathfrak{S}_i^k . The integrals J_i of \mathfrak{S}_i^0 over the domain Ω in the space $x_0 = \text{const.}$ are the *energy* ($i = 0$) and the *impulse* ($i = 1, 2, 3$) of the particle. If one integrates each of the four conservation laws:

$$(23) \quad \frac{\partial \mathfrak{S}_i^k}{\partial x_k} = 0,$$

which were proved above in general, in the same way then the first term ($k = 0$) will yield the temporal derivative dJ_i / dt . However, from **Gauss**’s theorem, the integrals of the other three terms will yield a “force flux” through the outer surface of Ω , which is expressed by an integral that is taken over that surface: viz., the components of the *field force* that acts upon the particle from the outside. The separation that comes about as a result of the splitting of space and time yields the juxtaposition of the *inertial force* dJ_i / dt and the field force that is characteristic of mechanics.

The integrand of the action principle (19), whose implications we shall now pursue, shall be called \mathfrak{B} . Since $\int \mathfrak{B} dx$ is not an invariant, the argument that was applied in Chap. II in order to prove the conservation laws cannot be maintained with no further assumptions. However, we also have $\delta' \int \mathfrak{B} dx = 0$ now for a variation δ that, from (14), is produced by an infinitely-small *displacement* in the true sense of the word: viz., $\pi = 0$, ξ^i constant. As far as that is concerned, we must make no assumptions about \mathfrak{B} whatsoever. If we set:

$$\delta \mathfrak{G} = \mathfrak{G}^{ik} \delta g_{ik} + \mathfrak{G}^{\alpha\beta i} \delta g_{\alpha\beta i}$$

then the formulas:

$$(24) \quad \frac{\partial (\bar{\mathfrak{S}}_i^k \xi^i)}{\partial x_k} = 0 \quad \text{with} \quad \bar{\mathfrak{S}}_i^k = \mathfrak{B} \delta_i^k + \frac{\partial g_{\alpha\beta}}{\partial x_i} \mathfrak{G}^{\alpha\beta k} + \lambda \frac{\partial \varphi^\alpha}{\partial x_i} \mathfrak{f}^{k\alpha}$$

will follow from this on the basis of the validity of **Hamilton**’s principle. However, these are not the conservation laws for energy and impulse. Moreover, in order to get them, we must next write down **Maxwell**’s equations in the form:

$$\frac{\partial \left(\pi \mathfrak{s}^k + \frac{\partial \pi}{\partial x_\alpha} \mathfrak{f}^{k\alpha} \right)}{\partial x_k} = 0,$$

in which we set $\pi = -(\varphi_i \xi^i)$ and add the equation that comes about, when multiplied by λ , to (24). Equations (23) then arise, and indeed, one will have:

$$\mathfrak{S}_i^k = \mathfrak{B} \delta_i^k + \frac{\partial g_{\alpha\beta}}{\partial x_i} \mathfrak{G}^{\alpha\beta,k} - \lambda f_{i\alpha} \mathfrak{f}^{k\alpha} - \lambda \varphi_i \mathfrak{s}^k.$$

This energy density is composed of three parts:

1. The term that is noticeable only in the interior of the material particle:

$$\lambda \left\{ \frac{1}{2} (\mathfrak{s}^r \varphi_r) \delta_i^k - \varphi_i \mathfrak{s}^k \right\}.$$

2. The one that belongs to **Maxwell's** field:

$$\lambda \left\{ i \delta_i^k - f_{i\alpha} \mathfrak{f}^{k\alpha} \right\}.$$

3. The gravitational energy:

$$\left(\frac{\alpha \sqrt{g}}{4} - \mathfrak{G} \right) \delta_i^k + \frac{\partial g_{\alpha\beta}}{\partial x_i} \mathfrak{G}^{\alpha\beta,k}.$$

We think of the range of values for the g_{ik} outside of the channel as being extended over the channel by constants when we “flatten out” and “bridge over” the fine, deep furrow that the path of the material particle digs in the metric face of the world and treat that current filament as a line in that flattened metric field. Let ds be the associated proper-time differential. We can introduce a coordinate system about a location along the current filament such that one has:

$$ds^2 = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2)$$

there, the direction of the current filament is given by:

$$dx_0 : dx_1 : dx_2 : dx_3 = 1 : 0 : 0 : 0,$$

and the derivatives $\partial g_{\alpha\beta} / \partial x_i$ vanish. For the cross-section $x_0 = \text{const.}$ of the current filament that one makes through that location, one will then have (approximately):

$$J_1 = J_2 = J_3 = 0,$$

as in the static case, assuming that the internal structure of the particle is the same as when it remained at rest in that coordinate system; that assumption is admissible for quasi-stationary acceleration. Likewise, of the integrals:

$$\int \mathfrak{s}^i dx_1 dx_2 dx_3$$

that are extended over the cross-section of the current filament, only the 0th one does not have the value 0 there, but is equal to the charge e of the particle (which is an invariant

that is independent of time, from the conservation law). Under such circumstances, the “force fluxes,” which represent the portion that arises from (3.), will drop out of the integral that are taken over the outer surface of the cap Ω at the moment considered. In order for that to be true, it is essential that the expressions (3.) depend not only linearly, but also quadratically, upon the differential quotients $\partial g_{\alpha\beta} / \partial x_i$. The part that arises from (1.) can be neglected, since $s^i = 0$ outside of the particle. Only (2.) remains, and that part yields the ponderomotive force of the electromagnetic field according to **Maxwell's** theory: viz., $e f_{0i}$ (f_{ik} is the external field here; the assertion is correct at least when that field does not vary too strongly in time relative to the particles). We get the equations:

$$\frac{dJ_i}{dt} = e f_{0i}.$$

If we revert to an arbitrary coordinate system then the following formulas will enter in place of the ones that were obtained:

$$J_i = m u_i, \quad \text{in which } u_i = \frac{dx_i}{ds}$$

and a proportionality factor m of “mass” will appear; furthermore:

$$(25) \quad \frac{d(mu_i)}{ds} - \frac{1}{2} \frac{\partial g_{\alpha\beta}}{\partial x_i} m u^\alpha u^\beta = m \cdot f_{ki} u^k.$$

Here, the g_{ik} , like the f_{ik} , refer to the flattened metric. The charge e is constant. If one multiplies the last equation by u_i and sums over i then one will find that:

$$\frac{dm}{ds} = 0,$$

so the mass will likewise be constant. It depends upon the choice of constant α in such a way that $m = \bar{m} \sqrt{\alpha}$ (\bar{m} is independent of α).

The connection with the ordinary formulas has been achieved. It is essential for their validity that the gauge is normalized by $F = \text{const.}$ For quasi-stationary acceleration, a clock will measure the integral $\int ds$ of the proper time that corresponds to that normalization. However, that result is linked with the action principle that is used as a basis here.

The problem of matter. – The fact that the conservation laws imply constant charge and mass for a material particle still does not explain the fact that all electrons possess the same charge and mass and consistently maintain them. The particles are never completely isolated from each other then since considerable deviations should not arise in the course of long periods of time. Moreover, that must rest upon the fact that the world-laws must admit only a discrete number of static solutions that would represent stable

corpuscles. With that, we come to true problem of matter: Can it be solved on the basis of the action principle that we have assumed here? It seems that the answer to that question might be “no,” since **Mie** has shown that the addition of a term to the **Maxwellian** action density (which is, by the way, a function of $q = \sqrt{\varphi_i \varphi^i}$) will certainly make matter impossible when that function does not vanish to at least fifth order for $q = 0$ ⁽¹⁾. However, for him, that knowledge arose from the fact that he required the regularity of the static spherically-symmetric solution at infinity. Here, however, those solutions will undoubtedly lead, not to an infinite space, but a closed one, so completely different regularity demands must be posed.

I must touch upon yet another point before I go on to explicit calculations. It is a fact that for the electron, pure numbers appear whose orders of magnitude are completely different from 1, such as the ratio of the electron radius to the gravitational radius of its mass, which has order to magnitude 10^{40} ; the ratio of the electron radius to the world-radius might have a similar order of magnitude. That would seem to demand that a pure number with an enormous value must be included in **Hamilton’s** principle from the outset, which is what happened with our Ansatz: viz., the constant β . On the other hand, one must then concede that the structure of the world should be based upon certain pure numbers with fortuitous numerical values that are abstract entities. A way around that dilemma is probably possible only when one assumes that a world-law does not prescribe a specific value for the number β , but demand only that it must be constant. In other words, it must read: Any virtual variation of the metric that vanishes outside of a finite world-domain for which $\delta \int l dx$ vanishes will also make the variation of:

$$\int \frac{1}{4} F^2 \sqrt{g} dx$$

vanish. In that way, the problem of matter will become an “eigenvalue” problem: Only certain discrete values of β will belong to regular solutions. They correspond to possible particles that nonetheless all exist in the same world, whether next to each other or inside of each other, which would impose reciprocal fine modifications of their internal structures. Noteworthy consequences for the organization of the universe seem to come to light then, along with the possibility of an explanation for the fact that it is globally at rest, but locally in a state of unrest.

In the static, spherically-symmetric case, we have two scalar fields c and φ , which depend upon only the distance $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, and the spatial line element $d\sigma^2$, which can, with the use of a suitable scale of distance, be conferred the form:

$$(dx_1^2 + dx_2^2 + dx_3^2) + p(x_1 dx_1 + x_2 dx_2 + x_3 dx_3)^2,$$

in which p is also a function that depends upon only r . I set:

⁽¹⁾ Ann. Phys. (Leipzig) **39** (1912), pp. 14.

$$w = h^2 = 1 + p r^2, \quad \Delta = \sqrt{g} = hc; \quad \frac{\varphi}{\Delta} = u, \quad \frac{r \varphi'}{\Delta} = v.$$

(The prime means differentiation with respect to r .) The spatial coordinate system is fixed up to a rotation by the normalization that has been performed, while the functions c and φ are fixed up to a common constant factor, and u , v , w are determined completely. The action principle (that one writes down with no further assumptions) implies the differential equations:

$$(26) \quad \left\{ \begin{array}{l} \Delta \Delta' = \frac{3}{4} w^2 \varphi^2 r, \\ \left(\frac{p r^3}{1 + p r^2} \right)' = \frac{\alpha r^2}{4} + \frac{3 \varphi^2 w r^2}{4 \Delta^2} - \frac{1}{2} \frac{\varphi'^2 r^2}{\Delta^2}, \\ \left(\frac{r^2 \varphi'}{\Delta} \right)' + \frac{3}{2\lambda} \frac{\varphi w r^2}{\Delta} = 0. \end{array} \right.$$

The problem is of order four and of such a type that the mathematician will hopefully sweep the sails before him. All the same, I can reduce order by one when I introduce the functions that were previously denoted by u , v , w . Instead of r , I shall employ the square $r^2 = \rho$ for a variable, and find that:

$$(D_v) \quad 2\rho \frac{dv}{d\rho} + v + \frac{3uw\rho}{2\lambda} = 0,$$

$$(D_w) \quad 2\rho \frac{dw}{d\rho} + w(w-1) - \frac{w^2}{4} (\alpha\rho + 3u^2 w\rho - 2\lambda v^2) = 0.$$

In addition, one has $r (u \Delta)' = v \Delta$; if I substitute the expression for Δ' / Δ that is implied by (26) in this then that will give:

$$(D_u) \quad 2\rho \frac{du}{d\rho} + \frac{3}{4} \rho u^3 w^2 - v = 0.$$

These differential equations (D) determine u , v , w ; one will get Δ from a quadrature of:

$$(27) \quad \frac{d \ln \Delta}{d\rho} = \frac{3}{8} (u w)^2.$$

If one prescribes the initial values: u arbitrary, $v = 0$, $w = 1$ then one will get power-series solutions that satisfy the equations formally. In the theory of differential equations, it is

shown that they converge ⁽¹⁾. We will then get ∞^1 solutions that are regular at the “pole” $\rho = 0$.

A solution that represents the evolution of the field in a material particle that is capable of existing will lead to a closed space. The equator of that space will be defined by $\rho = \rho_0$. In the vicinity of the equator, one must employ the quantity z that is introduced by way of:

$$\rho = \rho_0 (1 - z^2)$$

for the purpose of uniformization. w must then become infinite to order 2 for $z = 0$, while c and φ will remain regular, and c will certainly not vanish for $z = 0$. Δ will become infinite to order 1, while u and v will then take on zero loci of order 1 when $z = 0$. If I set:

$$\frac{u}{z} = \bar{u}, \quad \frac{v}{z} = \bar{v}, \quad w \cdot z^2 = \bar{w}$$

then \bar{u} , \bar{v} , \bar{w} will become regular, as well as even, functions of z . Let me point out that from (27), $\ln \Delta$ is a monotone-increasing function of ρ . The sign in this equation is fortunately arranged such that it allows a growth in Δ beyond all limits to be possible. If I employ $z^2 = t$ as the independent variable then that will produce the differential equations:

$$(\bar{D}) \quad \begin{cases} 2t \frac{d\bar{u}}{dt} + \bar{u} - \frac{3}{4} \rho_0 (1-t) \bar{u}^3 \bar{w}^2 - \frac{t \bar{v}}{1-t} = 0, \\ 2t \frac{d\bar{v}}{dt} + \frac{1-2t}{1-t} \bar{v} - \frac{3\rho_0}{2\lambda} \bar{u} \bar{w} = 0, \\ 2t \frac{d\bar{w}}{dt} - \frac{\bar{w}(\bar{w}-3t+2)}{1+t} + \frac{\bar{w}^2}{4} \left(\alpha \rho_0 + 3\rho_0 \bar{u}^2 \bar{w} - \frac{2\lambda t \bar{v}^2}{1-t} \right) = 0, \end{cases}$$

and

$$(28) \quad \frac{d \ln \Delta}{dt} = -\frac{3\rho_0}{8t} (\bar{u} \bar{w})^2.$$

By comparing the constant terms in the power series development, one will get the following initial values for $t = 0$:

$$\bar{u} = \frac{\alpha \rho_0 - 4}{2\sqrt{3\rho_0}}, \quad \bar{v} = \frac{\sqrt{3\rho_0}}{\lambda}, \quad \bar{w} = \frac{4}{\alpha \rho_0 - 4}.$$

As would follow from the existence theorem that was stated above, they are associated with a single regular solution of the system (\bar{D}) , along with a Δ that will become infinite like $1/\sqrt{t}$ [if the power-series development on the right-hand side of (28) begins with the term $-1/2t$]. Every value of ρ_0 will then correspond to a solution to the problem that

⁽¹⁾ **Picard**, *Traité d'Analyse*, 3, pp. 21.

is regular on the equator, and when one varies ρ_0 , one will get a family of ∞^1 such fields. Only those of them that belong to values:

$$\rho_0 > \frac{4}{\alpha} \quad \left(r_0 > \frac{2}{\sqrt{\alpha}} \right)$$

can come under consideration, since w must be positive; hence, the radius must have a cosmic magnitude! In the *three-dimensional* manifold of all solutions to the system of equations (D), we will then have the *one-dimensional* manifold of all fields that are regular at the pole and the *one-dimensional* manifold of all fields that are regular at the equator. These two manifolds will generally “intersect” as rarely as two lines in space. However, one should probably expect that there are isolated special values of λ – viz., *eigenvalues* – for which such an intersection comes about; i.e., a solution – viz., an “eigenfunction” – exists that remains regular at the pole, as well as at the equator. The present-day tools of analysis are hardly sufficient to prove the actual existence of those eigenvalues.

The putative world-law. – In Mie’s theory of gravitation, when regarded as an extension of Einstein’s, in the form that Hilbert presented it ⁽¹⁾, the **Hamilton** function $W (= \mathfrak{B} \sqrt{g})$ was constrained by only the demand that it had to be invariant under coordinate transformations. That demand allowed one considerable latitude. That latitude will be restricted rather sharply by our postulate that W must be an invariant of weight – 2 under changes of gauge, but still not to such a degree that W would be determined uniquely by it. If we assume that W is constructed rationally from the curvature components then, as far as I can see, that will suggest only the following five possibilities:

1. The **Maxwellian** one: $l = \frac{1}{4} f_{ik} f^{ik}$.
2. According to the same model, one can define $\frac{1}{4} F_{ik} F^{ik}$ from the vector curvature.

In this, multiplication is to be interpreted as the composition of matrices. The expression itself a matrix, but its trace L is a scalar of weight – 2:

$$L = \frac{1}{4} F_{\beta ik}^{\alpha} F_{\alpha}^{\beta ik}.$$

If $*L$ is the analog of the invariant that is defined by the direction curvature then one will have $L = *L + l$.

3. One permutes the indices β and i in the second factor $F_{\alpha}^{\beta ik}$ of the expression for L .

⁽¹⁾ **D. Hilbert**, Nachr. Ges. Wiss. zu Göttingen (1915), pp. 395.

4. The scalar $F_{ik} F^{ik}$ arises from the contracted tensor $F_{i\alpha k}^{\alpha} = F_{ik}$.
5. The invariant F^2 that was employed above.

The assertion that was made means that any invariant of the given kind can be composed from these five quantities linearly with constant coefficients.

The action principle that was implemented in the previous sections is constituted as follows: Its **Hamilton** function was a linear combination of (1.) and (5.). I believe that one can assert that this action principle implies everything that **Einstein's** theory has implied up to now, but in the more far-reaching questions of cosmology and the constitution of matter, it exhibits a clear superiority. Nevertheless, I do not believe that the laws of nature that are exactly applicable in reality are resolved by it. In regard to the actual character of the magnitude of the curvature, it seems to me that, in fact, the invariants (3.)-(5.) are artificial constructions, compared to the two natural ones, viz., the "principal invariants" (1.) and (3.). If I am not deceived by this faith in aesthetics (which correctly gives the four-dimensionality of the world) then the world-law would read: *Any virtual change in the metric that vanishes outside of a finite domain, and for which $\delta \int dx = 0$, will also fulfill the equation $\delta \int \mathcal{L} dx = 0$.* I think that I will pursue the consequences of that action principle in a continuation of this paper.

The fruitfulness of the new viewpoint of gauge invariance has shown itself, above all, in the problem of matter. However, the decisive conclusions in that context are fortified behind a wall of mathematical complexities that I have not been able to break through up to now.

(Received on 7 January 1919)
