SPECIAL THEORY OF RELATIVITY

by

C. W. KiLMISTER

Professor of Mathematics, King's College, London

PERGAMON PRESS

Oxford · New York

Toronto · Sydney · Braunschweig

NOTES ON EXTRACT 4

Poincaré starts by considering the experiments of Michelson and the explanation of Lorentz and Fitzgerald, and he has already in mind the later article of Lorentz (Extract 3). His own approach has been quite independent of Lorentz, and he remarks that the results which he has obtained agree in all important respects with those of Lorentz. The first section recapitulates the results already given by Lorentz, and he next proceeds, with typically French elegance, to deduce all of these results from a variational principle. This deduction is carried out in order that in the third section Poincaré can relate the invariance under the Lorentz transformation to the invariance of the variational principle. In the fourth section he goes on to show that the transformations do indeed form a group. Subsequent sections (which are omitted) are then concerned with rather technical matters of less interest now, but at the end of the paper there is a section in which Poincaré attempts to relate what he has found to the problem of gravitation. This problem is one which will occupy us exclusively in the succeeding book, since it found its complete solution not by means of field theories of the kind envisaged by Poincaré but by an entirely reformulated theory (general relativity) some years later It is, however, of the greatest interest to observe what an extremely sophisticated gravitational theory can be produced by Poincaré in a Lorentz invariant fashion.

One mathematical point needs to be noted. Poincaré adheres to the old usage of d for both partial and ordinary differentiation. He then introduces what would usually be a symbol of partial differentiation, with a specialised meaning, in Section 2.

EXTRACT 4[*]

The dynamics of the Electron

By H. Poincaré

Introduction

It would seem at first sight that the aberration of light and the optical and electrical effects related thereto should afford a means of determining the absolute motion of the Earth, or rather its motion relative to the ether instead of relative to the other celestial bodies. An attempt at this was made, indeed, by Fresnel, but he soon perceived that the Earth's motion does not affect the laws of refraction and reflection. Similar experiments, such as that using a waterfilled telescope, or any in which only the first-order terms relative to the aberration were considered, likewise yielded only negative results. The explanation of this was soon found; but Michelson, who devised an experiment wherein the terms involving the square of the aberration should be detectable, was equally unsuccessful.

This impossibility of experimentally demonstrating the absolute motion of the Earth appears to be a general law of Nature; it is reasonable to assume the existence of this law, which we shall call the *relativity postulate*, and to assume that it is universally valid. Whether this postulate, which so far is in agreement with experiment, be later confirmed or disproved by more accurate tests, it is, in any case, of interest to see what consequences follow from it.

[* *Rend. del Circ. Mat. di Palermo* **21**, 129–46 and 166–75 (1906).]

C. Kilmister ed 1970 [Poincaré 1906]

One explanation, suggested by Lorentz and Fitzgerald, involves the hypothesis that all bodies undergo a contraction in the direction of the Earth's motion, of an amount proportional to the square of the aberration; such a contraction, which we shall call the *Lorentz contraction*, would explain the result of Michelson's experiment and of all others conducted heretofore. The hypothesis would nevertheless be inadequate if the relativity postulate were valid in its most general form.

Lorentz has sought to extend and modify the hypothesis so as to make it fully compatible with the relativity postulate. This he has succeeded in doing, in his paper "Electromagnetic phenomena in a system moving with any velocity smaller than that of light" (*Proceedings of the Section of Sciences, Koninklijke Akademie van Wetenschappen te Amsterdam* **6**, 809–831, 1904).

In view of the importance of this problem, I resolved to examine it further. The results which I have obtained agree with those of Lorentz in all the principal points, and I have needed only to modify and augment them in certain details. These differences, which are of but minor importance, will be shown in later sections.

Lorentz's concept may be summarised thus: if a common translatory motion may be imparted to the entire system without any alteration of the observable phenomena, then the equations of an electromagnetic medium are unaltered by certain transformations, which we shall call *Lorentz transformations*. In this way two systems, of which one is fixed and the other is in translatory motion, become exact images of each other.

Langevin† sought to derive a modification of Lorentz's concept. Both authors consider that an electron in motion assumes the form of an oblate spheroid; but Lorentz considers that two of the axes of this spheroid remain constant, whereas Langevin supposes that its volume remains constant. These two authors

† Langevin had been anticipated by Bucherer of Bonn, who earlier put forward the same idea. See A. H. Bucherer, *Mathematische Einführung in die Elektronentheorie*, Teubner, Leipzig, 1904.

have shown that the two hypotheses are in agreement with the experiments of Kaufmann, as is Abraham's original hypothesis of a rigid spherical electron.

The advantage of Langevin's theory is that it involves only the electromagnetic forces and the constraints; but it is not compatible with the relativity postulate. This was shown by Lorentz, and I have likewise proved it by a different method, based upon the use of group theory.

We must return therefore to Lorentz's theory, but, in order to maintain this free from unacceptable contradictions, a special force must be invoked to account both for the contraction and for the constancy of two of the axes. I have attempted to determine this force, and have found that *it can be regarded as a constant external pressure acting upon an electron capable of deformation and compression, the work done being proportional to the change in the volume of the electron.*

Then, if the inertia of matter is exclusively of electromagnetic origin, as has been customarily supposed since Kaufmann's experiment, and if all forces (other than the constant pressure to which I have just alluded) are of electromagnetic origin, the relativity postulate can be accepted as strictly valid. I show this by means of a very simple calculation based upon the principle of least action.

But this is not all. Lorentz, in his paper already mentioned, has deemed it necessary to extend his hypothesis in such a manner that the postulate remains valid when there exist forces other than the electromagnetic forces. In Lorentz's view, all forces, no matter how originating, are affected by the Lorentz transformation (and therefore by a translatory motion) in the same manner as the electromagnetic forces.

It was necessary to consider this hypothesis more closely, and in particular to ascertain the changes which it would compel us to apply to the laws of gravitation.

First of all, we find that gravitational action would be propagated with the velocity of light, and not instantaneously. This might in itself appear to be sufficient reason to reject the hypothesis, for Laplace has shown that such propagation cannot occur. But, in fact, the effects of this are largely counterbalanced by another phenomenon, and there is, therefore, no contradiction between the proposed law and astronomical observations.

The question arises whether it is possible to discover a law which satisfies Lorentz's condition and which yet reduces to Newton's law whenever the velocities of the bodies are so small that the squares of these velocities (and the products of the accelerations and the distances) may be neglected in comparison with the square of the velocity of light.

It will be seen later than the answer must be affirmative.

Is the law, thus modified, compatible with astronomical observations?

At first sight it appears to be so, but a more detailed discussion is necessary to settle the question.

Even assuming, however, that the new hypothesis survives this test, what conclusion is to be drawn? If the gravitational attraction is propagated with the velocity of light, this cannot occur by mere chance, but must be dependent on the ether; we should then have to investigate the nature of this dependence, and attempt to relate it to other such dependences.

We cannot be satisfied with formulae that are merely placed side by side and agree only by a lucky chance; these formulae must, as it were, interlock. The mind will consent only when it sees the reason for the agreement, and when this agreement even seems to have been predictable.

But the matter may be viewed in a different light, as an analogy will show. Let us imagine some astronomer before Copernicus, pondering upon the Ptolemaic system. He would notice that, for every planet, either the epicycle or the deferent is traversed in the same time. This cannot be due to chance, and there must be some mysterious bond between all the planets of the system.

Then Copernicus, by a simple change of the co-ordinate axes which were supposed fixed, did away with this seeming relationship: every planet described one circular orbit only, and the periods of revolution became independent of one another—until Kepler once more established the relationship that had apparently been destroyed.

Now, there may be an analogy with our problem. If we assume the relativity postulate, we find a quantity common to the law of gravitation and the laws of electromagnetism, and this quantity is the velocity of light; and this same quantity appears in every other force, of whatever origin. There can be only two explanations.

Either, everything in the universe is of electromagnetic origin; or, this constituent which appears common to all the phenomena of physics has no real existence, but arises from our methods of measurement. What are these methods? One might first reply, the bringing into juxtaposition of objects regarded as invariable solid things; but this is no longer so in our present theory, if the Lorentz contraction is assumed. In this theory, two lengths are by definition equal if they are traversed by light in the same time.

Perhaps the abandonment of this definition would suffice to overthrow Lorentz's theory as decisively as the system of Ptolemy was by the work of Copernicus. Should this ever happen, it would by no means argue the futility of Lorentz's analysis: whatever the faults of the Ptolemaic theory, it was the necessary foundation for Copernicus to build upon.

I have therefore not hesitated to publish these incomplete results, even though at the present time the entire theory may seem to be threatened by the discovery of cathode rays.

§ 1. The Lorentz Transformation

Lorentz has adopted a particular system of units, such that the factors of 4π no longer appear in the formulae. I shall do likewise, and moreover I shall choose the units of length and of time in such a way that the velocity of light is equal to unity. Then, if f, g, h denote the electrical displacement; α, β, γ the magnetic force; F, G, H the vector potential; ψ the scalar potential; ϱ the electrical charge density; ξ, η, ζ the velocity of the electron; u, v, w the current, the fundamental equations become

$$\left.\begin{aligned}
u &= \frac{df}{dt}+\varrho\xi = \frac{d\gamma}{dy}-\frac{d\beta}{dz},\\
\alpha &= \frac{dH}{dy}-\frac{dG}{dz}, \quad f = -\frac{dF}{dt}-\frac{d\psi}{dx},\\
\frac{d\alpha}{dt} &= \frac{dg}{dz}-\frac{dh}{dy}, \quad \frac{d\varrho}{dt}+\Sigma\frac{d(\varrho\xi)}{dx} = 0,\\
\Sigma\frac{df}{dx} &= \varrho, \quad \frac{d\psi}{dt}+\Sigma\frac{dF}{dx} = 0,\\
\Box &= \Delta-\frac{d^2}{dt^2} = \Sigma\frac{d^2}{dx^2}-\frac{d^2}{dt^2},\\
\Box\psi &= -\varrho, \quad \Box F = -\varrho\xi.
\end{aligned}\right\} \tag{1}$$

An elementary particle of matter, having a volume $dx\,dy\,dz$, is acted upon by a mechanical force, whose components $X\,dx\,dy\,dz$, $Y\,dx\,dy\,dz$, $Z\,dx\,dy\,dz$ are given by the formula

$$X = \varrho f+\varrho(\eta\gamma-\zeta\beta). \tag{2}$$

These equations can be subjected to a remarkable transformation discovered by Lorentz, the significance of which is that it explains why no experimental demonstration of the absolute motion of the universe is possible. If we put

$$x' = kl(x+\varepsilon t), \quad t' = kl(t+\varepsilon x), \quad y' = ly, \quad z' = lz, \tag{3}$$

where l and ε are any constants, and

$$k = \frac{1}{\sqrt{(1-\varepsilon^2)}},$$

and if we also put

$$\Box' = \Sigma\frac{d^2}{dx'^2}-\frac{d^2}{dt'^2},$$

then

$$\Box' = \Box\,.\,l^{-2}.$$

Let a sphere be carried along with the electron in a uniform translatory motion, and let the equation of this moving sphere be

$$(x-\xi t)^2+(y-\eta t)^2+(z-\zeta t)^2 = r^2;$$

the volume of the sphere is then $\frac{4}{3}\pi r^3$.

The foregoing transformation will change the sphere into an ellipsoid, whose equation is easily found. From the equations (3), it immediately follows that

$$x = \frac{k}{l}(x'-\varepsilon t'), \quad t = \frac{k}{l}(t'-\varepsilon x'), \quad y = \frac{y'}{l}, \quad z = \frac{z'}{l}. \tag{3'}$$

The equation of the ellipsoid is then

$$k^2(x'-\varepsilon t'-\xi t'+\varepsilon\xi x')^2+(y'-\eta kt'+\eta k\varepsilon x')^2 +(z'-\zeta kt'+\zeta k\varepsilon x')^2 = l^2r^2.$$

The ellipsoid moves uniformly; when $t' = 0$, it is

$$k^2x'^2(1+\xi\varepsilon)^2+(y'+\eta k\varepsilon x')^2+(z'+\zeta k\varepsilon x')^2 = l^2r^2,$$

and its volume is

$$\frac{4}{3}\pi r^3\frac{l^3}{k(1+\xi\varepsilon)}.$$

If the charge on an electron is to be unaltered by the transformation, and if the new electrical charge density be denoted by

ϱ', it follows that

$$\varrho' = \frac{k}{l^3}(\varrho + \varepsilon\varrho\xi). \tag{4}$$

The new velocities ξ', η', ζ' will be given by

$$\xi' = \frac{dx'}{dt'} = \frac{d(x+\varepsilon t)}{d(t+\varepsilon x)} = \frac{\xi+\varepsilon}{1+\varepsilon\xi},$$

$$\eta' = \frac{dy'}{dt'} = \frac{dy}{k\,d(t+\varepsilon x)} = \frac{\eta}{k(1+\varepsilon\xi)},$$

$$\zeta' = \frac{\zeta}{k(1+\varepsilon\xi)},$$

whence

$$\varrho'\xi' = \frac{k}{l^3}(\varrho\xi+\varepsilon\varrho), \quad \varrho'\eta' = \frac{1}{l^3}\varrho\eta, \quad \varrho'\zeta' = \frac{1}{l^3}\varrho\zeta. \tag{4'}$$

Here I must for the first time indicate a disagreement with Lorentz's analysis. Lorentz (*op. cit.*, page 813, formulae (7) and (8)) writes, in our notation,

$$\varrho' = \frac{1}{kl^3}\varrho, \quad \xi' = k^2(\xi+\varepsilon), \quad \eta' = k\eta, \quad \zeta' = k\zeta.$$

These lead to the same relationships

$$\varrho'\xi' = \frac{k}{l^3}(\varrho\xi+\varepsilon\varrho), \quad \varrho'\eta' = \frac{1}{l^3}\varrho\eta, \quad \varrho'\zeta' = \frac{1}{l^3}\varrho\zeta,$$

but with a different value of ϱ'.

It should be noticed that formulae (4) and (4′) satisfy the continuity condition

$$\frac{d\varrho'}{dt} + \Sigma\frac{d(\varrho'\xi')}{dx'} = 0.$$

For, let λ be an undetermined coefficient, and D the Jacobian of

$$t+\lambda\varrho, \quad x+\lambda\varrho\xi, \quad y+\lambda\varrho\eta, \quad z+\lambda\varrho\zeta \tag{5}$$

with respect to t, x, y, z. Then

$$D = D_0 + D_1\lambda + D_2\lambda^2 + D_3\lambda^3 + D_4\lambda^4,$$

with $$D_0 = 1, \quad D_1 = \frac{d\varrho}{dt} + \Sigma\frac{d(\varrho\xi)}{dx} = 0.$$

Let $\lambda' = l^2\lambda$; then the four functions

$$t'+\lambda'\varrho', \quad x'+\lambda'\varrho'\xi', \quad y'+\lambda'\varrho'\eta', \quad z'+\lambda'\varrho'\zeta' \tag{5'}$$

are related to the functions (5) by the same linear relationships as those which exist between the old and new variables. If, therefore, D' denotes the Jacobian of the functions (5′) with respect to the new variables, then

$$D' = D, \quad D' = D_0' + D_1'\lambda' + \ldots + D_4'\lambda'^4,$$

whence

$$D_0' = D_0 = 1, \quad D_1' = l^{-2}D_1 = 0$$
$$= \frac{d\varrho'}{dt'} + \Sigma\frac{d(\varrho'\xi')}{dx'}, \qquad \text{q.e.d.}$$

With Lorentz's hypothesis, this condition would not be fulfilled, since the value of ϱ' is not the same.

The new vector and scalar potentials will be defined so as to satisfy the conditions

$$\Box'\psi' = -\varrho', \quad \Box'F' = -\varrho'\xi'. \tag{6}$$

Hence we find

$$\psi' = \frac{k}{l}(\psi+\varepsilon F), \quad F' = \frac{k}{l}(F+\varepsilon\psi), \quad G' = \frac{1}{l}G, \quad H' = \frac{1}{l}H. \tag{7}$$

These formulae are noticeably different from those of Lorentz, but the difference rests, ultimately, only on the definitions used.

The new electric and magnetic fields will be defined so as to satisfy the equations

$$f' = -\frac{dF'}{dt'} - \frac{d\psi'}{dx'}, \quad \alpha' = \frac{dH'}{dy'} - \frac{dG'}{dz'}. \tag{8}$$

$\lambda' = l^4\lambda$

$l^{-4}D_1$

It is easily seen that

$$\frac{d}{dt'} = \frac{k}{l}\left(\frac{d}{dt} - \varepsilon\frac{d}{dx}\right), \quad \frac{d}{dx'} = \frac{k}{l}\left(\frac{d}{dx} - \varepsilon\frac{d}{dt}\right),$$

$$\frac{d}{dy'} = \frac{1}{l}\frac{d}{dy}, \quad \frac{d}{dz'} = \frac{1}{l}\frac{d}{dz},$$

and therefore

$$\left.\begin{aligned} f' &= \frac{1}{l^2}f, \quad g' = \frac{k}{l^2}(g+\varepsilon\gamma), \quad h' = \frac{k}{l^2}(h-\varepsilon\beta), \\ \alpha' &= \frac{1}{l^2}\alpha, \quad \beta' = \frac{k}{l^2}(\beta-\varepsilon h), \quad \gamma' = \frac{k}{l^2}(\gamma+\varepsilon g). \end{aligned}\right\} \quad (9)$$

These formulae are identical with those of Lorentz.

Our transformation does not affect equations (1): the continuity condition and equations (6) and (8) are identical with some of the equations (1) if the primes are omitted.

The equations (6), together with the continuity condition, give

$$\frac{d\psi'}{dt'} + \Sigma\frac{dF'}{dx'} = 0. \quad (10)$$

We have only to prove that

$$\frac{df'}{dt'} + \varrho'\xi' = \frac{d\gamma'}{dy'} - \frac{d\beta'}{dz'}, \quad \frac{d\alpha'}{dt'} = \frac{dg'}{dz'} - \frac{dh'}{dy'},$$

$$\Sigma\frac{df'}{dx'} = \varrho',$$

and it is easily seen that these relationships necessarily follow from equations (6), (8) and (10).

Let us now make a comparison of the forces before and after the transformation.

Let X, Y, Z be the force before the transformation, and X', Y', Z' the force after it, both per unit volume. If X' is to satisfy the same equations as before the transformation, we must have

$$X' = \varrho' f' + \varrho'(\eta'\gamma' - \zeta'\beta'),$$
$$Y' = \varrho' g' + \varrho'(\zeta'\alpha' - \xi'\gamma'),$$
$$Z' = \varrho' h' + \varrho'(\xi'\beta' - \eta'\alpha'),$$

or, substituting the expressions (4), (4′) and (9) and using equations (2),

$$\left.\begin{aligned} X' &= \frac{k}{l^5}(X + \varepsilon\Sigma X\xi), \\ Y' &= \frac{1}{l^5}Y, \\ Z' &= \frac{1}{l^5}Z. \end{aligned}\right\} \quad (11)$$

If X_1, Y_1, Z_1 denote the components of the force per unit electric charge on the electron, and X_1', Y_1', Z_1' the same quantities after the transformation, then

$$X_1 = f + \eta\gamma - \zeta\beta, \quad X_1' = f' + \eta'\gamma' - \zeta'\beta',$$
$$X = \varrho X_1, \quad X' = \varrho' X_1',$$

and we should obtain

$$\left.\begin{aligned} X_1' &= \frac{k}{l^5}\frac{\varrho}{\varrho'}(X_1 + \varepsilon\Sigma X_1\xi), \\ Y_1' &= \frac{1}{l^5}\frac{\varrho}{\varrho'}Y_1, \\ Z_1' &= \frac{1}{l^5}\frac{\varrho}{\varrho'}Z_1. \end{aligned}\right\} \quad (11')$$

Lorentz's result was, in our notation (*op. cit.*, page 813, formula (10)),

$$\left.\begin{aligned} X_1 &= l^2 X_1' - l^2\varepsilon(\eta' g' + \zeta' h'), \\ Y_1 &= \frac{l^2}{k}Y_1' + \frac{l^2\varepsilon}{k}\xi' g', \\ Z_1 &= \frac{l^2}{k}Z_1' + \frac{l^2\varepsilon}{k}\xi' h'. \end{aligned}\right\} \quad (11'')$$

Before proceeding, it is necessary to ascertain the reason for this considerable difference. It occurs, evidently, because the formulae for ξ', η', ζ' are not the same, whereas those for the electric and magnetic fields are the same.

If the inertia of the electrons is of purely electromagnetic origin, and if moreover they are subject only to forces of electromagnetic origin, the condition of equilibrium requires that, within the electrons,

$$X = Y = Z = 0.$$

From the relations (11), these are clearly equivalent to

$$X' = Y' = Z' = 0.$$

Thus the equilibrium conditions are unaffected by the transformation.

Unfortunately, such a simple hypothesis is inadmissible. For, if we assume that $\xi = \eta = \zeta = 0$, the conditions $X = Y = Z = 0$ will imply that $f = g = h = 0$, and therefore

$$\Sigma \frac{df}{dx} = 0, \quad \text{i.e. } \varrho = 0.$$

Similar results would be obtained in the general case. Hence we must assume that there are not only electromagnetic forces but also either other forces or constraints. We then have to determine the conditions governing these forces or constraints such that the equilibrium of the electrons is unaffected by the transformation. This will be done in a subsequent section.

§ 2. The Principle of Least Action

Lorentz's derivation of his equations from the principle of least action is well known. I shall, however, discuss this point further (although I have nothing essential to add to Lorentz's analysis), since I prefer to present it in a slightly different form, which will

be of use later. I write

$$J = \int dt\, d\tau [\tfrac{1}{2}\Sigma f^2 + \tfrac{1}{2}\Sigma \alpha^2 - \Sigma F u], \tag{1}$$

with f, α, F, u, etc., assumed subject to the following conditions and those obtained from them by symmetry:

$$\Sigma \frac{df}{dx} = \varrho, \quad \alpha = \frac{dH}{dy} - \frac{dG}{dz}, \quad u = \frac{df}{dt} + \varrho\xi. \tag{2}$$

The integral J is taken over the following ranges:

(a) the whole of space, for the volume element $d\tau = dx\, dy\, dz$;
(b) the interval between t_0 and t_1, for the time element dt.

According to the principle of least action, the integral J must have a minimum value when the quantities in it satisfy:

(a) the conditions (2);
(b) the condition that the system is in specified states at the limiting times t_0 and t_1.

The latter condition enables us to transform the integrals, using an integration by parts with respect to the time. For, given an integral of the form

$$\int dt\, d\tau\, A \frac{dB\, \delta C}{dt},$$

where C is one of the quantities defining the state of the system, and δC the variation of C, integration by parts with respect to the time shows that this integral is equal to

$$\int d\tau \Big[AB\, \delta C \Big]_{t=t_0}^{t=t_1} - \int dt\, d\tau \frac{dA}{dt} dB\, \delta C.$$

Since the state of the system at the limiting times is specified, $\delta C = 0$ for $t = t_0$ and for $t = t_1$; the first integral is therefore zero, and only the second integral remains.

We can effect a similar integration by parts with respect to x, y or z, since

$$\int A\frac{dB}{dx}\,dx\,dy\,dz\,dt = \int AB\,dy\,dz\,dt - \int B\frac{dA}{dx}\,dx\,dy\,dz\,dt.$$

The integrations extend to infinity, and in the first integral on the right-hand side we must therefore put $x = \pm\infty$; this integral is then zero, because all the functions are assumed to tend to zero at infinity, and we have

$$\int A\frac{dB}{dx}\,d\tau\,dt = -\int B\frac{dA}{dx}\,d\tau\,dt.$$

If the system were assumed subject to constraints, the constraint conditions would have to be included among the conditions to be satisfied by the various quantities appearing in the integral J.

First, let F, G, H receive increments δF, δG, δH; then

$$\delta\alpha = \frac{d\,\delta H}{dy} - \frac{d\,\delta G}{dz}.$$

We must have

$$\delta J = \int dt\,d\tau\left[\Sigma\alpha\left(\frac{d\,\delta H}{dy} - \frac{d\,\delta G}{dz}\right) - \Sigma u\,\delta F\right] = 0,$$

or, on integrating by parts,

$$\delta J = \int dt\,d\tau\left[\Sigma\left(\delta G\frac{\delta\alpha}{\delta z} - \delta H\frac{d\alpha}{dy}\right) - \Sigma u\,\delta F\right]$$

$$= -\int dt\,d\tau\Sigma\,\delta F\left(u - \frac{d\gamma}{dy} + \frac{d\beta}{dz}\right) = 0,$$

whence, equating to zero the coefficient of the arbitrary quantity δF,

$$u = \frac{d\gamma}{dy} - \frac{d\beta}{dz}. \qquad (3)$$

From this we obtain (using an integration by parts)

$$\int \Sigma Fu\,d\tau = \int \Sigma F\left(\frac{d\gamma}{dy} - \frac{d\beta}{dz}\right)d\tau$$

$$= \int \Sigma\left(\beta\frac{dF}{dz} - \gamma\frac{dF}{dy}\right)d\tau$$

$$= \int \Sigma\alpha\left(\frac{dH}{dy} - \frac{dG}{dz}\right)d\tau,$$

or

$$\int \Sigma Fu\,d\tau = \int \Sigma\alpha^2\,d\tau,$$

whence finally

$$J = \int dt\,d\tau\left(\tfrac{1}{2}\Sigma f^2 - \tfrac{1}{2}\Sigma\alpha^2\right). \qquad (4)$$

Henceforward, having regard to the relation (3), δJ is independent of δF, and therefore of $\delta\alpha$. Let us now vary the other quantities.

The expression (1) for J gives

$$\delta J = \int dt\,d\tau(\Sigma f\,\delta f - \Sigma F\,\delta u).$$

But f, g, h must satisfy the first condition (2), so that

$$\Sigma\frac{d\,\delta f}{dx} = \delta\varrho, \qquad (5)$$

and we may write

$$\delta J = \int dt\,d\tau\left[\Sigma f\,\delta f - \Sigma F\,\delta u - \psi\left(\Sigma\frac{d\,\delta f}{dx} - \delta\varrho\right)\right]. \qquad (6)$$

From the calculus of variations, it is known that the calculation should be made as if ψ were an arbitrary function, δJ were represented by the expression (6), and the variations were not subject to the condition (5).

We also have

$$\delta u = \frac{d\,\delta f}{dt} + \delta(\varrho\xi),$$

and therefore, on integration by parts,

$$\delta J = \int dt\, d\tau \Sigma\, \delta f\left(f + \frac{dF}{dt} + \frac{d\psi}{dx}\right) + \int dt\, d\tau(\psi\, \delta\varrho - \Sigma F\, \delta(\varrho\xi)). \quad (7)$$

If now it be assumed that the electrons undergo no variation, then $\delta\varrho = \delta(\varrho\xi) = 0$, and the second integral vanishes. For δJ to be zero, we must have

$$f + \frac{dF}{dt} + \frac{d\psi}{dx} = 0. \quad (8)$$

In the general case, therefore,

$$\delta J = \int dt\, d\tau(\psi\, \delta\varrho - \Sigma F\, \delta(\varrho\xi)). \quad (9)$$

It remains to determine the forces acting upon the electrons. To do so, we must assume that a complementary force $-X\,d\tau$, $-Y\,d\tau$, $-Z\,d\tau$ is applied to each electron volume element, and write down the condition for this force to balance the forces of electromagnetic origin. Let U, V, W be the components of the displacement of the electron volume element $d\tau$, measured from any given initial position. Let δU, δV, δW be the variations of this displacement. The virtual work corresponding to the complementary force will be

$$-\int \Sigma X\, \delta U\, d\tau,$$

and the equilibrium condition just mentioned will therefore be

$$\delta J = -\int \Sigma X\, \delta U\, d\tau\, dt. \quad (10)$$

In order to transform δJ, we first seek the equation of continuity stating that the electron charge remains constant under the variation.

Let x_0, y_0, z_0 be the initial position of the electron. Its position at the time considered will be

$$x = x_0 + U, \quad y = y_0 + V, \quad z = z_0 + W.$$

We shall define also an auxiliary variable ε to generate the variation of each function: for any function A,

$$\delta A = \delta\varepsilon \frac{dA}{d\varepsilon}.$$

This is done because it will be convenient to be able to change between the notation of the calculus of variations and that of the ordinary differential calculus whenever desired.

The functions under consideration may be regarded in two ways: (a) as functions of the five variables x, y, z, t, ε, so that the position remains unaltered when only t and ε vary, in which case derivatives with be denoted by d as usual; (b) as functions of the five variables $x_0, y_0, z_0, t, \varepsilon$, so that a particular electron is followed when only t and ε vary, in which case derivatives will be denoted by the symbol ∂. Then we have

$$\xi = \frac{\partial U}{\partial t} = \frac{\partial U}{\partial t} + \xi\frac{dU}{dx} + \eta\frac{dU}{dy} + \zeta\frac{dU}{dz} = \frac{\partial x}{\partial t}. \quad (11)$$

Now, let Δ denote the Jacobian of x, y, z with respect to x_0, y_0, z_0:

$$\Delta = \frac{\partial(x, y, z)}{\partial(x_0, y_0, z_0)}.$$

If t receives an increment ∂t while $\varepsilon, x_0, y_0, z_0$ remain constant, there will be consequent increments $\partial x, \partial y, \partial z$ of x, y, z, and $\partial\Delta$ of Δ, with

$$\partial x = \xi\,\partial t, \quad \partial y = \eta\,\partial t, \quad \partial z = \zeta\,\partial t,$$

$$\Delta + \partial\Delta = \frac{\partial(x+\partial x, y+\partial y, z+\partial z)}{\partial(x_0, y_0, z_0)}$$

whence

$$1+\frac{\partial \Delta}{\Delta} = \frac{\partial(x+\partial x,\, y+\partial y,\, z+\partial z)}{\partial(x,\, y,\, z)}$$

$$= \frac{\partial(x+\xi\, \partial t,\, y+\eta\, \partial t,\, z+\zeta\, \partial t)}{\partial(x,\, y,\, z)}.$$

From this we obtain

$$\frac{1}{\Delta}\frac{\partial \Delta}{\partial t} = \frac{d\xi}{dx}+\frac{d\eta}{dy}+\frac{d\zeta}{dz}. \qquad (12)$$

Since the mass of an electron is constant,

$$\frac{\partial(\varrho\Delta)}{\partial t} = 0, \qquad (13)$$

and therefore

$$\frac{\partial\varrho}{\partial t}+\Sigma\varrho\frac{d\xi}{dx} = 0, \qquad \frac{\partial\varrho}{\partial t} = \frac{d\varrho}{dt}+\Sigma\xi\frac{d\varrho}{dx},$$

$$\frac{d\varrho}{dt}+\Sigma\frac{d(\varrho\xi)}{dx} = 0.$$

These are the various forms of the equation of continuity with respect to the variable t. Similar forms can be deduced with respect to the variable ε. Let

$$\delta U = \frac{\partial U}{\partial\varepsilon}\delta\varepsilon, \qquad \delta V = \frac{\partial V}{\partial\varepsilon}\delta\varepsilon, \qquad \delta W = \frac{\partial W}{\partial\varepsilon}\delta\varepsilon;$$

then

$$\delta U = \frac{dU}{d\varepsilon}\delta\varepsilon+\delta U\frac{dU}{dx}+\delta V\frac{dU}{dy}+\delta W\frac{dU}{dz}, \qquad (11')$$

$$\frac{1}{\Delta}\frac{\partial\Delta}{\partial\varepsilon} = \Sigma\frac{\partial U}{\partial\varepsilon}, \qquad \frac{\partial(\varrho\Delta)}{\partial\varepsilon} = 0, \qquad (12')$$

$$\left.\begin{aligned}\delta\varepsilon\frac{\partial\varrho}{\partial\varepsilon}+\Sigma\varrho\frac{d\delta U}{dx} = 0, \qquad \frac{\partial\varrho}{\partial\varepsilon} = \frac{d\varrho}{d\varepsilon}+\Sigma\frac{\partial U}{\partial\varepsilon}\frac{d\varrho}{dx},\\ \delta\varrho+\Sigma\frac{d\varrho\,\delta U}{dx} = 0.\end{aligned}\right\} \qquad (13')$$

It will be seen that there is a difference between the definition of $\delta U = (\partial U/\partial\varepsilon)\delta\varepsilon$ and that of $\delta\varrho = (d\varrho/d\varepsilon)\delta\varepsilon$, and that this definition of δU is the one which is appropriate to the formula (10).

The first term in equation (9) can be transformed by means of the last equation (13′):

$$\int dt\, d\tau\psi\, \delta\varrho = -\int dt\, d\tau\psi\Sigma\frac{d\varrho\, dU}{dx},$$

or, after integration by parts,

$$\int dt\, d\tau\psi\, \delta\varrho = \int dt\, d\tau\, \Sigma\varrho\frac{d\psi}{dx}\delta U. \qquad (14)$$

Let us now seek to determine

$$\delta(\varrho\xi) = \frac{d(\varrho\xi)}{d\varepsilon}\delta\varepsilon.$$

We may notice that $\varrho\Delta$ can depend only on x_0, y_0, z_0; for, if an electron volume element be considered whose initial position is a rectangular parallelepiped with edges dx_0, dy_0, dz_0, the charge on this element is

$$\varrho\Delta\, dx_0\, dy_0\, dz_0.$$

Since the charge must remain constant,

$$\frac{\partial(\varrho\Delta)}{\partial t} = \frac{\partial(\varrho\Delta)}{\partial\varepsilon} = 0. \qquad (15)$$

Hence we have

$$\frac{\partial^2(\varrho\,\Delta U)}{\partial t\,\partial\varepsilon} = \frac{\partial}{\partial\varepsilon}\left(\varrho\Delta\frac{\partial U}{\partial t}\right) = \frac{\partial}{\partial t}\left(\varrho\Delta\frac{\partial U}{\partial\varepsilon}\right). \qquad (16)$$

For any function A we have, by the equation of continuity,

$$\frac{1}{\Delta}\frac{\partial(A\Delta)}{\partial t} = \frac{dA}{dt}+\Sigma\frac{d(A\xi)}{dx}$$

and similarly

$$\frac{1}{\Delta}\frac{\partial(A\Delta)}{\partial\varepsilon} = \frac{dA}{d\varepsilon} + \Sigma\frac{d(A\,\partial U/\partial\varepsilon)}{dx}.$$

Hence

$$\frac{1}{\Delta}\frac{\partial}{\partial\varepsilon}\left(\varrho\Delta\frac{\partial U}{\partial t}\right) = \frac{d(\varrho\,\partial U/\partial t)}{d\varepsilon} + \frac{d(\varrho(\partial U/\partial t)(\partial U/\partial\varepsilon))}{dx}$$
$$+\frac{d(\varrho(\partial U/\partial t)(\partial V/\partial\varepsilon))}{dy} + \frac{d(\varrho(\partial U/\partial t)(\partial W/\partial\varepsilon))}{dz}, \quad (17)$$

$$\frac{1}{\Delta}\frac{\partial}{\partial t}\left(\varrho\Delta\frac{\partial U}{\partial\varepsilon}\right) = \frac{d(\varrho\,\partial U/\partial\varepsilon)}{dt} + \frac{d(\varrho(\partial U/\partial t)(\partial U/\partial\varepsilon))}{dx}$$
$$+\frac{d(\varrho(\partial V/\partial t)(\partial U/\partial\varepsilon))}{dy} + \frac{d(\varrho(\partial W/\partial t)(\partial U/\partial\varepsilon))}{dz}. \quad (17')$$

The right-hand sides of (17) and (17′) must be equal, and, since

$$\frac{\partial U}{\partial t} = \xi, \quad \frac{\partial U}{\partial\varepsilon}\delta\varepsilon = \delta U, \quad \frac{d(\varrho\xi)}{d\varepsilon}\delta\varepsilon = \delta(\varrho\xi),$$

we obtain

$$\delta(\varrho\xi) + \frac{d(\varrho\xi\,\delta U)}{dx} + \frac{d(\varrho\xi\,\delta V)}{dy} + \frac{d(\varrho\xi\,\delta W)}{dz}$$
$$= \frac{d(\varrho\,\delta U)}{dt} + \frac{d(\varrho\xi\,\delta U)}{dx} + \frac{d(\varrho\eta\,\delta U)}{dy} + \frac{d(\varrho\zeta\,\delta U)}{dz}. \quad (18)$$

Now transforming the second term in (9), we have

$$\int dt\,d\tau\,\Sigma F\,\delta(\varrho\xi)$$
$$= \int dt\,d\tau\left[\Sigma F\frac{d(\varrho\,\delta U)}{dt} + \Sigma F\frac{d(\varrho\eta\,\delta U)}{dy} + \Sigma F\frac{d(\varrho\zeta\,\delta U)}{dz}\right.$$
$$\left. - \Sigma F\frac{d(\varrho\xi\,\delta V)}{dy} - \Sigma F\frac{d(\varrho\xi\,\delta W)}{dz}\right].$$

Integration by parts on the right-hand side yields

$$\int dt\,d\tau\left[-\Sigma\varrho\,\delta U\frac{dF}{dt} - \Sigma\varrho\eta\,\delta U\frac{dF}{dy} - \Sigma\varrho\zeta\,\delta U\frac{dF}{dz}\right.$$
$$\left. + \Sigma\varrho\xi\,\delta V\frac{dF}{dy} + \Sigma\varrho\xi\,\delta W\frac{dF}{dz}\right].$$

Next we note that

$$\Sigma\varrho\xi\,\delta V\frac{dF}{dy} = \Sigma\varrho\zeta\,\delta U\frac{dH}{dx},$$
$$\Sigma\varrho\xi\,\delta W\frac{dF}{dz} = \Sigma\varrho\eta\,\delta U\frac{dG}{dx}.$$

For, if the sums on either side are expanded, they become identities. Since also

$$\frac{dH}{dx} - \frac{dF}{dz} = -\beta, \quad \frac{dG}{dx} - \frac{dF}{dy} = \gamma,$$

the right-hand side becomes

$$\int dt\,d\tau\left[-\Sigma\varrho\,\delta U\frac{dF}{dt} + \Sigma\varrho\gamma\eta\,\delta U - \Sigma\varrho\beta\zeta\,\delta U\right],$$

and thus finally

$$\delta J = \int dt\,d\tau\,\Sigma\varrho\,\delta U\left(\frac{d\psi}{dx} + \frac{dF}{dt} + \beta\zeta - \gamma\eta\right)$$
$$= \int dt\,d\tau\,\Sigma\varrho\,\delta U(-f + \beta\zeta - \gamma\eta).$$

Equating the coefficients of δU on either side of (10), we have

$$X = f - \beta\zeta + \gamma\eta.$$

This is equation (2) of Section 1.

§ 3. The Lorentz Transformation and the Principle of Least Action

Let us consider whether the principle of least action can explain the success of the Lorentz transformation. First of all, we must examine the result of applying this transformation to the integral

$$J = \int dt\, d\tau(\tfrac{1}{2}\Sigma f^2 - \tfrac{1}{2}\Sigma\alpha^2)$$

(formula (4) of section 2).

We have firstly

$$dt'\, d\tau' = l^4\, dt\, d\tau,$$

since x', y', z', t' are related to x, y, z, t by linear expressions whose determinant is l^4. Next,

$$\left.\begin{aligned} l^4\Sigma f'^2 &= f^2+k^2(g^2+h^2)+k^2\varepsilon^2(\beta^2+\gamma^2)+2k^2\varepsilon(g\gamma-h\beta), \\ l^4\Sigma\alpha'^2 &= \alpha^2+k^2(\beta^2+\gamma^2)+k^2\varepsilon^2(g^2+h^2)+2k^2\varepsilon(g\gamma-h\beta) \end{aligned}\right\} \quad (1)$$

(formulae (9) of section 1), whence

$$l^4(\Sigma f'^2 - \Sigma\alpha'^2) = \Sigma f^2 - \Sigma\alpha^2.$$

Thus, if we put

$$J' = \int dt'\, d\tau'(\tfrac{1}{2}\Sigma f'^2 - \tfrac{1}{2}\Sigma\alpha'^2),$$

the result is

$$J' = J.$$

However, for this equation to be valid, the limits of integration must be the same. Hitherto we have assumed that t ranged from t_0 to t_1, and x, y, z from $-\infty$ to $+\infty$. The limits of integration would then be altered by the Lorentz transformation; but there is no bar to assuming that $t_0 = -\infty$, $t_1 = +\infty$, and the limits for J and for J' are then the same.

We have thus to compare the two following equations, which are analogues of equation (10) in section 2:

$$\left.\begin{aligned} \delta J &= -\int \Sigma X\, \delta U\, d\tau\, dt, \\ \delta J' &= -\int \Sigma X'\, \delta U'\, d\tau'\, dt'. \end{aligned}\right\} \quad (2)$$

To do so, we must first compare $\delta U'$ with δU.

Let us consider an electron having initial co-ordinates x_0, y_0, z_0. Its co-ordinates at the instant t will be

$$x = x_0+U, \quad y = y_0+V, \quad z = z_0+W.$$

If the corresponding electron after the Lorentz transformation is considered, its co-ordinates will be

$$x' = kl(x+\varepsilon t), \quad y' = ly, \quad z' = lz,$$

where

$$x' = x_0+U', \quad y' = y_0+V', \quad z' = z_0+W';$$

but these values will be reached at the instant

$$t' = kl(t+\varepsilon x).$$

If the variables are subjected to variations δU, δV, δW, while at the same time t receives an increment δt, then the total increments of the co-ordinates x, y, z will be

$$\delta x = \delta U+\xi\, \delta t, \quad \delta y = \delta V+\eta\, \delta t, \quad \delta z = \delta W+\zeta\, \delta t.$$

Similarly,

$$\delta x' = \delta U'+\xi'\, \delta t', \; \delta y' = \delta V'+\eta'\, \delta t', \; \delta z' = \delta W'+\zeta'\, \delta t',$$

and, by the Lorentz transformation,

$$\delta x' = kl(\delta x+\varepsilon\, \delta t), \quad \delta y' = l\, \delta y, \quad \delta z' = l\, \delta z,$$
$$\delta t' = kl(\delta t+\varepsilon\, \delta x);$$

hence, assuming $\delta t = 0$, we find

$$\delta x' = \delta U' + \xi' \, \delta t' = kl \, \delta U,$$
$$\delta y' = \delta V' + \eta' \, \delta t' = l \, \delta V,$$
$$\delta t' = kl\varepsilon \, \delta U.$$

Since

$$\xi' = \frac{\xi+\varepsilon}{1+\xi\varepsilon}, \quad \eta' = \frac{\eta}{k(1+\xi\varepsilon)},$$

we have, on replacing $\delta t'$ by its value,

$$kl(1+\xi\varepsilon) \, \delta U = \delta U'(1+\xi\varepsilon) + (\xi+\varepsilon)kl\varepsilon \, \delta U,$$
$$l(1+\xi\varepsilon) \, \delta V = \delta V'(1+\xi\varepsilon) + \eta l\varepsilon \, \delta U.$$

Using the definition of k, we obtain from these equations

$$\delta U = \frac{k}{l} \, \delta U' + \frac{k\varepsilon}{l} \xi \, \delta U',$$

$$\delta V = \frac{1}{l} \, \delta V' + \frac{k\varepsilon}{l} \eta \, \delta U',$$

and similarly

$$\delta W = \frac{1}{l} \, \delta W' + \frac{k\varepsilon}{l} \zeta \, \delta U';$$

hence

$$\Sigma X \, \delta U = \frac{1}{l}(kX \, \delta U' + Y \, \delta V' + Z \, \delta W') + \frac{k\varepsilon}{l} \, \delta U' \Sigma X\xi. \qquad (3)$$

Now, according to the equations (2), we must have

$$\int \Sigma X' \, \delta U' \, dt' \, d\tau' = \int \Sigma X \, \delta U \, dt \, d\tau = \frac{1}{l^4} \int \Sigma X \, \delta U \, dt' \, d\tau'.$$

Replacing $\Sigma X \, \delta U$ by its value (3) and equating coefficients, we find

$$X' = \frac{k}{l^5} X + \frac{k\varepsilon}{l^5} \Sigma X\xi, \quad Y' = \frac{1}{l^5} Y, \quad Z' = \frac{1}{l^5} Z.$$

These are the equations (11) of Section 1. Thus the principle of least action leads to the same results as does the analysis given in Section 1.

Returning to formulae (1), we see that $\Sigma f^2 - \Sigma \alpha^2$ is unaltered by the Lorentz transformation, apart from a constant factor. The same is not true of the expression $\Sigma f^2 + \Sigma \alpha^2$ which appears in the energy. If we consider only the case where ε is so small that its square may be neglected, so that $k = 1$, and if we also assume that $l = 1$, then

$$\Sigma f'^2 = \Sigma f^2 + 2\varepsilon(g\gamma - h\beta),$$
$$\Sigma \alpha'^2 = \Sigma \alpha^2 + 2\varepsilon(g\gamma - h\beta),$$

and, by addition,

$$\Sigma f'^2 + \Sigma \alpha'^2 = \Sigma f^2 + \Sigma \alpha^2 + 4\varepsilon(g\gamma - h\beta).$$

§ 4. The Lorentz Group

It is noteworthy that the Lorentz transformations form a group. For, if we put

$$x' = kl(x+\varepsilon t), \quad y' = ly, \quad z' = lz, \quad t' = kl(t+\varepsilon x),$$

and

$$x'' = k'l'(x'+\varepsilon' t'), \quad y'' = l'y', \quad z'' = l'z', \quad t'' = k'l'(t'+\varepsilon' x'),$$

with

$$k^{-2} = 1-\varepsilon^2, \quad k'^{-2} = 1-\varepsilon'^2,$$

we find that

$$x'' = k''l''(x+\varepsilon'' t), \quad y'' = l''y,$$
$$z'' = l''z, \quad t'' = k''l''(t+\varepsilon'' x),$$

with

$$\varepsilon'' = \frac{\varepsilon+\varepsilon'}{1+\varepsilon\varepsilon'}, \quad l'' = ll', \quad k'' = kk'(1+\varepsilon\varepsilon') = \frac{1}{\sqrt{(1-\varepsilon''^2)}}.$$

Taking $l = 1$ and assuming ε infinitesimal, with

$$x' = x+\delta x, \quad y' = y+\delta y, \quad z' = z+\delta z, \quad t' = t+\delta t,$$

we have

$$\delta x = \varepsilon t, \quad \delta y = \delta z = 0, \quad \delta t = \varepsilon x.$$

This is the infinitesimal generating transformation of the group, which I shall denote by T_1, and which in Lie's notation may be written

$$t\frac{d\phi}{dx}+x\frac{d\phi}{dt} = T_1.$$

If we take $\varepsilon = 0$ and $l = 1+\delta l$, on the other hand, we obtain

$$\delta x = x\ \delta l, \quad \delta y = y\ \delta l, \quad \delta z = z\ \delta l, \quad \delta t = t\ \delta l,$$

which yields another infinitesimal transformation T_0 of the group (assuming that l and ε are regarded as independent variables); in Lie's notation,

$$T_0 = x\frac{d\phi}{dx}+y\frac{d\phi}{dy}+z\frac{d\phi}{dz}+t\frac{d\phi}{dt}.$$

It is also possible to assign to the y-axis or to the z-axis the particular significance which has been given to the x-axis, thus obtaining two further infinitesimal transformations

$$T_2 = t\frac{d\phi}{dy}+y\frac{d\phi}{dt},$$

$$T_3 = t\frac{d\phi}{dz}+z\frac{d\phi}{dt},$$

which likewise would leave Lorentz's equations unchanged.

The combinations defined by Lie, such as

$$[T_1, T_2] = x\frac{d\phi}{dy}-y\frac{d\phi}{dx},$$

can also be constructed; but it is easily seen that this transformation is equivalent to a rotation of the co-ordinate axes through a very small angle about the z-axis. It is therefore not surprising that

this does not affect the form of Lorentz's equations, which are obviously independent of the axes chosen.

We are thus led to consider a continuous group, to be called the *Lorentz group*, possessing the following infinitesimal transformations:

(1) the transformation T_0, which commutes with every other;
(2) the three transformations T_1, T_2, T_3;
(3) the three rotations $[T_1, T_2]$, $[T_2, T_3]$, $[T_3, T_1]$.

Any transformation belonging to this group can be resolved into a transformation having the form

$$x' = lx, \quad y' = ly, \quad z' = lz, \quad t' = lt$$

and a linear transformation which leaves unaltered the quadratic form

$$x^2+y^2+z^2-t^2.$$

The group can also be generated in another way. Any transformation of the group may be regarded as comprising a transformation having the form

$$x' = kl(x+\varepsilon t), \quad y' = ly, \quad z' = lz, \quad t' = kl(t+\varepsilon x), \quad (1)$$

preceded and followed by an appropriate rotation.

For our purposes, however, we have to consider only certain of the transformations in this group. We must regard l as being a function of ε, the function being chosen so that this partial group, which will be denoted by P, is itself a group.

Let the system be rotated through 180° about the y-axis; then the resulting transformation must also belong to P. This operation is equivalent to changing the signs of x, x', z and z'; hence we have

$$x' = kl(x-\varepsilon t), \quad y' = ly, \quad z' = lz, \quad t' = kl(t-\varepsilon x). \quad (2)$$

Thus l is unchanged when ε is replaced by $-\varepsilon$.

Next, if P is a group, the substitution inverse to (1), which is

$$x' = \frac{k}{l}(x-\varepsilon t), \quad y' = \frac{y}{l}, \quad z' = \frac{z}{l}, \quad t' = \frac{k}{l}(t-\varepsilon x), \qquad (3)$$

must likewise belong to P; it must therefore be identical with (2), so that

$$l = 1/l.$$

Consequently, we must have $l = 1$.

[Note: there follow here four sections of a technical character, dealing with details of the electron theory not now of importance. Then Poincaré continues:]

§ 9. Hypotheses Concerning Gravitation

Thus Lorentz' theory would entirely account for the impossibility of demonstrating absolute motion, provided that all forces were of electromagnetic origin.

But there exist forces, such as gravitation, which cannot be regarded as being of electromagnetic origin. It may happen that two systems of bodies create equivalent electromagnetic fields, in the sense of exerting the same action upon electrified bodies and currents, while at the same time these two systems do not exert the same gravitational action upon Newtonian masses. The gravitational field is therefore not identical with the electromagnetic field. Lorentz was thus compelled to augment his hypothesis by assuming that *forces, of whatever origin, and in particular gravitation, are affected by translation* (or, if one prefers, by the Lorentz transformation) *in the same way as the electromagnetic forces.*

We must now examine this hypothesis in detail. If the Newtonian force is to behave in such a way under the Lorentz transformation, we can no longer suppose that this force depends only on the relative position of the attracting and the attracted body at the instant concerned; it must depend also on the velocities of

the two bodies. Moreover, we may reasonably assume that the force acting upon the attracted body, at an instant t, depends on the position and velocity of the body at that instant; but it will also depend on the position and velocity of the *attracting* body, not at the instant t but at some *previous* instant, as if gravitation required a certain time for its propagation.

Let us consider therefore the position of the attracted body at the instant t_0, and let its co-ordinates at that instant be x_0, y_0, z_0, and the components of its velocity be ξ, η, ζ; and let us consider the attracting body at the corresponding instant t_0+t, its co-ordinates at that instant being x_0+x, y_0+y, z_0+z, and its velocity components ξ_1, η_1, ζ_1.

First of all, we must have a relationship

$$\phi(t, x, y, z, \xi, \eta, \zeta, \xi_1, \eta_1, \zeta_1) = 0 \qquad (1)$$

to determine the time t. This relationship expresses the law of propagation of gravitational action; I shall by no means impose the condition that propagation occurs with the same velocity in every direction.

Next, let X_1, Y_1, Z_1 be the three components of the action exerted upon the attracted body at the instant t. We have to express X_1, Y_1, Z_1 as functions of SCHWARTZ HAS t_0

$$t, x, y, z, \xi, \eta, \zeta, \xi_1, \eta_1, \zeta_1. \qquad (2)$$

The conditions to be satisfied are as follows.

1. The relationship (1) must not be affected by the transformations of the Lorentz group.

2. The components X_1, Y_1, Z_1 must behave, under the Lorentz transformations, in the same manner as the electromagnetic forces denoted by the same letters, that is, as shown by equations (11′) of Section 1.

3. When both bodies are at rest, the usual law of attraction must apply.

In the latter case, however, it should be noted that the relationship (1) plays no part, since the time t is of no significance if both bodies are at rest.

The problem thus stated is clearly indeterminate. We shall therefore seek to satisfy as many further conditions as possible.

4. Astronomical observations do not appear to reveal any perceptible deviation from Newton's law, and we shall therefore choose the solution which differs least from this law when the velocities of the two bodies are small.

5. We shall attempt to ensure that t is always negative; for, whereas it is reasonable that the effect of gravitation should require a certain time for its propagation, we should find it more difficult to understand how this effect could depend on a position of the attracting body which the latter has *not yet reached*.

There is one case where the problem is no longer indeterminate, namely if the two bodies are at *relative* rest, i.e. if

$$\xi = \xi_1, \quad \eta = \eta_1, \quad \zeta = \zeta_1;$$

we shall therefore first investigate this case, assuming that these velocities are constant, and therefore that the two bodies are executing a common uniform motion of translation in a straight line.

We may assume that the x-axis has been taken to be parallel to this motion of translation, so that $\eta = \zeta = 0$, and we shall take $\varepsilon = -\xi$.

If, under these conditions, we apply the Lorentz transformation, the two bodies will be at rest after the transformation, with

$$\xi' = \eta' = \zeta' = 0$$

The components X_1', Y_1', Z_1' must then be in accordance with Newton's law and we have, apart from a constant factor,

$$\left.\begin{aligned} X_1' = -\frac{x'}{r'^3}, \quad Y_1' = -\frac{y'}{r'^3}, \quad Z_1' = -\frac{z'}{r'^3}, \\ r'^2 = x'^2+y'^2+z'^2. \end{aligned}\right\} \tag{3}$$

But, from Section 1,

$$x' = k(x+\varepsilon t), \quad y' = y, \quad z' = z, \quad t' = k(t+\varepsilon x),$$

$$\frac{\varrho'}{\varrho} = k(1+\xi\varepsilon) = k(1-\varepsilon^2) = \frac{1}{k}, \quad \Sigma X_1\xi = -X_1\varepsilon,$$

$$X_1' = k\frac{\varrho}{\varrho'}(X_1+\varepsilon\Sigma X_1\xi) = k^2X_1(1-\varepsilon^2) = X_1,$$

$$Y_1' = \frac{\varrho}{\varrho'}Y_1 = kY_1,$$

$$Z_1' = kZ_1.$$

Moreover,

$$x+\varepsilon t = x-\xi t, \quad r'^2 = k^2(x-\xi t)^2+y^2+z^2,$$

and

$$X_1 = \frac{-k(x-\xi t)}{r'^3}, \quad Y_1 = \frac{-y}{kr'^3}, \quad Z_1 = \frac{-z}{kr'^3}, \tag{4}$$

which may also be written

$$X_1 = \frac{dV}{dx}, \quad Y_1 = \frac{dV}{dy}, \quad Z_1 = \frac{dV}{dz}; \quad V = \frac{1}{kr'}. \tag{4'}$$

It seems at first sight that the indeterminacy remains, since no hypotheses have been made concerning the value of t, that is, concerning the velocity of propagation. Moreover, x is a function of t. But it is easily seen that the quantities $x-\xi t$, y and z which appear in the formulae do not depend on t.

Thus, if the two bodies have a common translatory motion, the force acting upon the attracted body is normal to an ellipsoid having the attracting body at its centre.

In order to proceed further, it is necessary to ascertain the *invariants of the Lorentz group*.

It is known that the substitutions forming this group (if $l = 1$) are linear and such that the quadratic form

$$x^2+y^2+z^2-t^2$$

is invariant. Putting

$$\xi = \frac{\delta x}{\delta t}, \qquad \eta = \frac{\delta y}{\delta t}, \qquad \zeta = \frac{\delta z}{\delta t};$$

$$\xi_1 = \frac{\delta_1 x}{\delta_1 t}, \quad \eta_1 = \frac{\delta_1 y}{\delta_1 t}, \quad \zeta_1 = \frac{\delta_1 z}{\delta_1 t},$$

we see that the Lorentz transformation causes δx, δy, δz, δt and $\delta_1 x$, $\delta_1 y$, $\delta_1 z$, $\delta_1 t$ to undergo the same linear substitutions as x, y, z, t.

If

$$\begin{matrix} x & y & z & t\sqrt{-1} \\ \delta x & \delta y & \delta z & \delta t\sqrt{-1} \\ \delta_1 x & \delta_1 y & \delta_1 z & \delta_1 t\sqrt{-1} \end{matrix}$$

are regarded as the co-ordinates of three points P, P', P'' in four-dimensional space, we see that the Lorentz transformation is simply a rotation of this space about a fixed origin. The only distinct invariants are therefore the six distances of the points P, P', P'' from one another and from the origin, or alternatively the two expressions

$$x^2+y^2+z^2-t^2, \qquad x\,\delta x+y\,\delta y+z\,\delta z-t\,\delta t$$

and the four expressions of the same form obtained by permuting the three points P, P', P'' in any manner.

What we are seeking, however, is invariant functions of the ten variables (2); we must therefore find, among combinations of the six invariants, those which depend only on these ten variables, i.e. those which are homogeneous and of degree zero with respect to δx, δy, δz, δt and with respect to $\delta_1 x$, $\delta_1 y$, $\delta_1 z$, $\delta_1 t$. This leaves four distinct invariants, namely

$$\Sigma x^2-t^2, \quad \frac{t-\Sigma x\xi}{\sqrt{(1-\Sigma\xi^2)}}, \quad \frac{t-\Sigma x\xi_1}{\sqrt{(1-\Sigma\xi_1^2)}}, \quad \frac{1-\Sigma\xi\xi_1}{\sqrt{[(1-\Sigma\xi^2)\,(1-\xi_1^2)]}}. \tag{5}$$

Let us now consider how the components of the force are transformed. We return to equations (11) of Section 1, which refer not to the force X_1, Y_1, Z_1 discussed here but to the force X, Y, Z per unit volume. Putting

$$T = \Sigma X\xi,$$

we see that these equations (11) may be written (with $l = 1$)

$$\left.\begin{aligned} X' &= k(X+\varepsilon T), & T' &= k(T+\varepsilon X), \\ Y' &= Y, & Z' &= Z; \end{aligned}\right\} \tag{6}$$

thus, X, Y, Z, T are transformed in the same manner as x, y, z, t. The invariants of the group will therefore be

$$\Sigma X^2-T^2, \quad \Sigma Xx-Tt, \quad \Sigma X\,\delta x-T\,\delta t, \quad \Sigma X\,\delta_1 x-T\,\delta_1 t.$$

The quantities in which we are interested are not X, Y, Z, but X_1, Y_1, Z_1, with

$$T_1 = \Sigma X_1\xi.$$

Evidently

$$\frac{X_1}{X} = \frac{Y_1}{Y} = \frac{Z_1}{Z} = \frac{T_1}{T} = \frac{1}{\varrho}.$$

Thus the Lorentz transformation will act upon X_1, Y_1, Z_1, T_1 in the same way as upon X, Y, Z, T, except that these expressions will in addition be multiplied by

$$\frac{\varrho}{\varrho'} = \frac{1}{k(1+\xi\varepsilon)} = \frac{\delta t}{\delta t'}.$$

Likewise, the transformation will act upon $\xi, \eta, \zeta, 1$ in the same way as upon $\delta x, \delta y, \delta z, \delta t$, except that these expressions will in addition be multiplied by the *same* factor,

$$\frac{\delta t}{\delta t'} = \frac{1}{k(1+\xi\varepsilon)}.$$

Let us now regard $X, Y, Z, T\sqrt{-1}$ as being the co-ordinates of a fourth point Q; the invariants will then be functions of the distances between the five points

$$O, P, P', P'', Q;$$

and these functions must be homogeneous of degree zero, firstly with respect to

$$X, Y, Z, T, \quad \delta x, \quad \delta y, \quad \delta z, \quad \delta t$$

(which variables can subsequently be replaced by $X_1, Y_1, Z_1, T_1, \xi, \eta, \zeta, 1$), and secondly with respect to

$$\delta_1 x, \quad \delta_1 y, \quad \delta_1 z, \quad 1$$

(which variables can subsequently be replaced by $\xi_1, \eta_1, \zeta_1, 1$).

In this way we find, in addition to the four invariants (5), four further and distinct invariants, namely

$$\frac{\Sigma X_1^2 - T_1^2}{1-\Sigma\xi^2}, \quad \frac{\Sigma X_1 x - T_1 t}{\sqrt{(1-\Sigma\xi^2)}}, \quad \frac{\Sigma X_1\xi_1 - T_1}{\sqrt{[(1-\Sigma\xi^2)(1-\Sigma\xi_1^2)]}}, \quad \frac{\Sigma X_1\xi - T_1}{1-\Sigma\xi^2}. \tag{7}$$

The last of these is always zero, according to the definition of T_1.

Which are the conditions that must now be satisfied?

1. The left-hand side of equation (1), which defines the velocity of propagation, must be a function of the four invariants (5).

It is obvious that a large number of hypotheses could be constructed. We shall consider only two of these.

(A) It may be that

$$\Sigma x^2 - t^2 = r^2 - t^2 = 0,$$

whence $t = \pm r$; and, since t must be negative, $t = -r$. This means that the velocity of propagation is equal to that of light. At first sight, it seems that this hypothesis should be rejected immediately; for Laplace has shown that the propagation is either instantaneous or much more rapid than that of light. But Laplace was discussing the hypothesis of a finite velocity of propagation alone, whereas here it is compounded with many others, and there may happen to be some more or less complete mutual compensation between them, a situation of which many examples have already appeared in the applications of the Lorentz transformation.

(B) It may be that

$$\frac{t-\Sigma x\xi_1}{\sqrt{(1-\Sigma\xi_1^2)}} = 0, \quad t = \Sigma x\xi_1.$$

The velocity of propagation is then much more rapid than that of light, but in certain cases t might be negative, which, as we have said, seems hardly acceptable. *We shall therefore abide by hypothesis* (A).

misprint in original: positive [correction by Schwartz] n10, 1287

2. The four invariants (7) must be functions of the invariants (5).

3. When both bodies are at absolute rest, X_1, Y_1, Z_1 must have the values given by Newton's law; when the bodies are at relative rest, the values must be those given by equations (4).

In the case of absolute rest, the first two invariants (7) must reduce to

$$\Sigma X_1^2, \quad \Sigma X_1 x,$$

or, by Newton's law, to

$$1/r^4, \quad -1/r.$$

According to hypothesis (A), the second and third of the invariants (5) become

$$\frac{-r-\Sigma x\xi}{\sqrt{(1-\Sigma\xi^2)}}, \quad \frac{-r-\Sigma x\xi_1}{\sqrt{(1-\Sigma\xi_1^2)}},$$

that is, for absolute rest,

$$-r, \quad -r.$$

We may therefore assume, *for example*, that the first two invariants (5) reduce to

$$\frac{(1-\Sigma\xi_1^2)^2}{(r+\Sigma x\xi_1)^4} - \frac{\sqrt{(1-\Sigma\xi_1^2)}}{r+\Sigma x\xi_1};$$

but other combinations are possible.

It is necessary to choose some combination, and a third equation is also needed in order to determine X_1, Y_1, Z_1. In making the choice, we shall attempt to remain as close as possible to Newton's law. Let us then examine the result when the squares of the velocities ξ, η, etc., are neglected (and $t = -r$). The four invariants (5) then become

$$0, \quad -r-\Sigma x\xi, \quad -r-\Sigma x\xi_1, \quad 1,$$

and the four invariants (7) become

$$\Sigma X_1^2, \quad \Sigma X_1(x+\xi r), \quad \Sigma X_1(\xi_1-\xi), \quad 0.$$

In order to compare this with Newton's law, however, a further transformation is necessary. In these equations, x_0+x, y_0+y, z_0+z represent the co-ordinates of the attracting body at the instant t_0+t, and $r = \sqrt{\Sigma x^2}$; in Newton's law, we have to consider the co-ordinates x_0+x_1, y_0+y_1, z_0+z_1 of the attracting body at the instant t_0, and the distance $r_1 = \sqrt{\Sigma x_1^2}$.

We may neglect the square of the time t occupied by the propagation, and therefore regard the motion as uniform; then

$$x = x_1+\xi_1 t, \quad y = y_1+\eta_1 t, \quad z = z_1+\zeta_1 t,$$
$$r(r-r_1) = \Sigma x\xi_1 t;$$

or, since $t = -r$,

$$x = x_1-\xi_1 r, \quad y = y_1-\eta_1 r, \quad z = z_1-\zeta_1 r, \quad r = r_1-\Sigma x\xi_1,$$

and the four invariants (5) become

$$0, \quad -r_1+\Sigma x(\xi_1-\xi), \quad -r_1, \quad 1$$

and the four invariants (7)

$$\Sigma X_1^2, \quad \Sigma X_1[x_1+(\xi-\xi_1)r_1], \quad \Sigma X_1(\xi_1-\xi), \quad 0.$$

In the second of these expressions I have written r_1 in place of r, since r is multiplied by $\xi-\xi_1$ and the square of ξ is neglected.

Newton's law gives, for these four invariants (7),

$$\frac{1}{r_1^4}, \quad -\frac{1}{r_1}-\frac{\Sigma x_1(\xi-\xi_1)}{r_1^2}, \quad \frac{\Sigma x_1(\xi-\xi_1)}{r_1^3}, \quad 0.$$

If therefore we denote the second and third invariants (5) by A and B, and the first three invariants (7) by M, N and P, Newton's law will be obeyed, to within terms of the order of the squares of the velocities, by putting

$$M = \frac{1}{B^4}, \quad N = \frac{+A}{B^2}, \quad P = \frac{A-B}{B^3}. \tag{8}$$

This solution is not unique: if the fourth invariant (5) is denoted by C, then $C-1$ is of the order of ξ^2, as is $(A-B)^2$. We may therefore add to the right-hand side of each of the equations (8) a term consisting of $C-1$ multiplied by any function of A, B and C, and a term consisting of $(A-B)^2$ also multiplied by any function of A, B and C.

The solution (8) appears the simplest at first sight, but it cannot be accepted. Since M, N and P are functions of X_1, Y_1, Z_1 and $T_1 = \Sigma X_1\xi$, these equations yield values of X_1, Y_1 and Z_1; but the resulting values may in some cases be imaginary.

In order to avoid this difficulty, we proceed differently, putting

$$k_0 = \frac{1}{\sqrt{(1-\Sigma\xi^2)}}, \quad k_1 = \frac{1}{\sqrt{(1-\Sigma\xi_1^2)}},$$

by analogy with

$$k = \frac{1}{\sqrt{(1-\varepsilon^2)}},$$

as in the Lorentz substitution.

Then, with the condition $-r = t$, the invariants (5) become

$$0, \quad A = -k_0(r+\Sigma x\xi), \quad B = -k_1(r+\Sigma x\xi_1),$$
$$C = k_0k_1(1-\Sigma\xi\xi_1).$$

Moreover, the following systems of quantities:

$$\begin{array}{llll} x, & y, & z, & -r = t \\ k_0X_1, & k_0Y_1, & k_0Z_1, & k_0T_1 \\ k_0\xi, & k_0\eta, & k_0\zeta, & k_0 \\ k_1\xi_1, & k_1\eta_1, & k_1\zeta_1, & k_1 \end{array}$$

are seen to undergo the *same* linear substitutions when the transformations of the Lorentz group are applied to them. We therefore put

$$\left.\begin{aligned} X_1 &= x\frac{\alpha}{k_0}+\xi\beta+\xi_1\frac{k_1}{k_0}\gamma, \\ Y_1 &= y\frac{\alpha}{k_0}+\eta\beta+\eta_1\frac{k_1}{k_0}\gamma, \\ Z_1 &= z\frac{\alpha}{k_0}+\zeta\beta+\zeta_1\frac{k_1}{k_0}\gamma, \\ T_1 &= -r\frac{\alpha}{k_0}+\beta+\frac{k_1}{k_0}\gamma. \end{aligned}\right\} \qquad (9)$$

It is evident that, if α, β, γ are invariants, X_1, Y_1, Z_1, T_1 will satisfy the fundamental condition, i.e. will undergo an appropriate linear substitution when the Lorentz transformations are applied to them.

If the equations (9) are compatible, we must have

$$\Sigma X_1\xi - T_1 = 0.$$

When X_1, Y_1, Z_1, T_1 are replaced by their values (9), the result is, after multiplication by k_0^2,

$$-A\alpha-\beta-C\gamma = 0. \qquad (10)$$

The desired conclusion is that the values of X_1, Y_1, Z_1 should remain in accordance with Newton's law when the squares of the velocities ξ, etc., and the products of the accelerations and the distances are neglected in comparison with the square of the velocity of light.

We can take

$$\beta = 0, \quad \gamma = -A\alpha/C.$$

To the approximation used,

$$k_0 = k_1 = 1, \quad C = 1, \quad A = -r_1+\Sigma x(\xi_1-\xi),$$
$$B = -r_1, \quad x = x_1+\xi_1 t = x_1-\xi_1 r.$$

Then the first equation (9) becomes

$$X_1 = \alpha(x-A\xi_1).$$

But, if ξ^2 is neglected, $A\xi_1$ may be replaced by $-r_1\xi_1$, or by $-r\xi_1$, whence

$$X_1 = \alpha(x+\xi_1 r) = \alpha x_1.$$

Newton's law would give

$$X_1 = -x_1/r_1^3.$$

We must therefore take as the invariant α one which reduces to $-1/r_1^3$ within the approximation adopted, that is, $1/B^3$. The equations (9) then become

$$\left.\begin{aligned} X_1 &= \frac{x}{k_0B^3}-\xi_1\frac{k_1}{k_0}\frac{A}{B^3C}, \\ Y_1 &= \frac{y}{k_0B^3}-\eta_1\frac{k_1}{k_0}\frac{A}{B^3C}, \\ Z_1 &= \frac{z}{k_0B^3}-\zeta_1\frac{k_1}{k_0}\frac{A}{B^3C}, \\ T_1 &= -\frac{r}{k_0B^3}-\frac{k_1}{k_0}\frac{A}{B^3C}. \end{aligned}\right\} \qquad (11)$$

sign?

It is seen, first of all, that the corrected attraction consists of two components, one parallel to the vector joining the positions of the two bodies, and the other parallel to the velocity of the attracting body.

When we speak of the position or the velocity of the attracting body, we mean its position or velocity at the instant when the gravitational wave leaves it; but the position or the velocity of the attracted body means its position or velocity at the instant when the gravitational wave reaches it, this wave being assumed to be propagated with the velocity of light.

I believe that it would be premature to attempt to continue the discussion of these formulae, and I shall therefore confine myself to making a few comments.

1. The solutions (11) are not unique; for the common factor $1/B^3$ may be replaced by

$$\frac{1}{B^3}+(C-1)f_1(A, B, C)+(A-B)^2 f_2(A, B, C),$$

where f_1 and f_2 are any functions of A, B and C. Moreover, β need not be taken as zero; any additional terms may be added to α, β and γ which satisfy the condition (10) and are of the second order n ξ for α, and of the first order in ξ for β and γ.

2. The first equation (11) may be written

$$X_1 = \frac{k_1}{B^3 C}\,[x(1-\Sigma\xi\xi_1)+\xi_1(r+\Sigma x\xi)], \tag{11'}$$

and the quantity in the brackets may in turn be written

$$(x+r\xi_1)+\eta(\xi_1 y-x\eta_1)+\zeta(\xi_1 z-x\zeta_1), \tag{12}$$

so that the total force is divisible into three components corresponding to the three parentheses in equation (12). The first component is somewhat similar to the mechanical force due to the electric field, the other two to the mechanical force due to the magnetic field. By virtue of comment 1, I may replace $1/B^3$ in equations (11) by C/B^3, so that X_1, Y_1, Z_1 are linear functions of the velocity ξ, η, ζ of the attracted body, C having been eliminated from the denominator of (11′). This completes the analogy.

Putting then

$$\left.\begin{aligned} k_1(x+r\xi_1) &= \lambda, & k_1(y+r\eta_1) &= \mu, & k_1(z+r\zeta_1) &= \nu,\\ k_1(\eta_1 z-\zeta_1 y) &= \lambda', & k_1(\zeta_1 x-\xi_1 z) &= \mu', & k_1(\xi_1 y-x\eta_1) &= \nu', \end{aligned}\right\} \tag{13}$$

with C eliminated from the denominator of (11′) we obtain

$$\left.\begin{aligned} X_1 &= \frac{\lambda}{B^3}+\frac{\eta\nu'-\zeta\mu'}{B^3},\\ Y_1 &= \frac{\mu}{B^3}+\frac{\zeta\lambda'-\xi\nu'}{B^3},\\ Z_1 &= \frac{\nu}{B^3}+\frac{\xi\mu'-\eta\lambda'}{B^3}, \end{aligned}\right\} \tag{14}$$

and also

$$B^2 = \Sigma\lambda^2-\Sigma\lambda'^2. \tag{15}$$

Thus λ, μ, ν or λ/B^3, μ/B^3, ν/B^3 is a kind of electric field, while λ', μ', ν' or λ'/B^3, μ'/B^3, ν'/B^3 is a kind of magnetic field.

3. The relativity postulate would compel us to use either the solution (11) or the solution (14) or any one of the solutions obtained therefrom by using comment 1. But the prime question is whether these are compatible with astronomical observations. The deviation from Newton's law is of the order of ξ^2, that is, 10,000 times less than if it had been of the order of ξ, as it would have been with the velocity of propagation equal to that of light and the other conditions unchanged. We may therefore hope that the deviation will not be very great; but only a more extended investigation will furnish the answer to this question.

Paris

July 1905