American State Government

AMERICAN STATE GOVERNMENT

By

W. BROOKE GRAVES

CHIEF OF THE STATE LAW SECTION
LEGISLATIVE REFERENCE SERVICE
LIBRARY OF CONGRESS

THIRD EDITION

D. C. HEATH AND COMPANY : Boston

THE THEORY OF RELATIVITY

BY

C. MØLLER

PROFESSOR OF MATHEMATICAL PHYSICS
IN THE UNIVERSITY OF COPENHAGEN

OXFORD
AT THE CLARENDON PRESS

Oxford University Press, Amen House, London E.C. 4
GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI CAPE TOWN IBADAN
Geoffrey Cumberlege, Publisher to the University

FIRST EDITION 1952

REPRINTED LITHOGRAPHICALLY IN GREAT BRITAIN
AT THE UNIVERSITY PRESS, OXFORD
FROM CORRECTED SHEETS OF THE FIRST EDITION
1955

PREFACE

THE present monograph is a somewhat extended version of a course of lectures which I have given at the University of Copenhagen during the last twenty years. Consequently, it is primarily a textbook for students in physics whose mathematical and physical training does not go beyond the methods of non-relativistic mechanics and electrodynamics. The intention has been to give an account of what may be called the classical theory of relativity in which all quantum effects are disregarded. In view of the paramount importance of quantum phenomena in modern physics, the limitation of the subject to classical phenomena might be considered a serious defect of the book. However, there are several important reasons for such a limitation of the subject. At present, a complete self-consistent relativistic quantum theory does not exist. Moreover, the classical theory of relativity, which by itself gives an admirably precise description of a very extended field of physical phenomena, must be the starting-point for the future development of a consistent relativistic quantum theory. For a student and research worker in this field an intimate aquaintance with the principles and methods of the classical theory of relativity is, therefore, just as indispensable as is the knowledge of the methods of Newtonian mechanics for a real understanding of ordinary quantum mechanics. Apart from this, the classical theory of relativity is one of the most fascinating and beautiful parts of theoretical physics on account of its inner consistency and the simplicity and generality of its basic assumptions.

The presentation of the subject in the present volume differs somewhat from the usual one in that the four-dimensional formulation of the theory plays a less dominant role than in most of the current textbooks. Certainly the four-dimensional representation, which is based on the symmetry between the space and time variables revealed by the discovery of the Lorentz transformation, is the most elegant way of expressing the principle of relativity in mathematical language, and it has been of the utmost importance for the rapid development, particularly of the general theory of relativity. In the early books on relativity it was, therefore, quite natural to emphasize as strongly as possible this newly discovered similarity between the space and time variables. However, in a textbook of today I think it is useful to stress again the fundamental physical difference between space and time, which was somewhat concealed by the purely formal four-dimensional representation.

In the first three chapters we have, therefore, avoided any reference to the four-dimensional picture, and the kinematics and point mechanics of the special theory of relativity are fully developed by means of the usual three-dimensional vector calculus. But in the following chapters also, where the elegant methods of the four-dimensional tensor calculus are developed and applied, a three-dimensional formulation, which gives a better insight into the physical meaning of the theory, is frequently given. As an example, I shall mention the treatment in §§ 110 and 111 of a freely falling particle in a given gravitational field. The motion of the particle is, of course, completely described by the statement that the time track of the particle in 4-space is a geodesic line, but this quasi-geometrical description does not convey a physical understanding of the phenomenon. In the three-dimensional physical space, however, the motion of the particle can be described by an equation of motion of the same type as that for a particle subject to an arbitrary force in a system of inertia, the only difference being that the geometry in the physical space is in general non-Euclidean. In this way we obtain definite expressions for the gravitational force on the particle as well as for the mass, momentum, and total energy of a particle moving with arbitrary velocity in a given gravitational field. The three-dimensional point of view thus leads to a reintroduction of dynamical concepts into the gravitational theory, which, I believe, makes it easier for the student fully to grasp the physical content of the general theory of relativity.

Since a real understanding of a physical theory is possible only through an intimate knowledge of its predecessors, the whole of Chapter I has been devoted to an historical survey of the difficulties of the non-relativistic theories. Many students who intend to specialize in experimental physics may feel that the time and effort which are needed to learn the methods of the general tensor calculus are out of proportion to the use they can make of this formalism in their future work. Such readers will find the main results of the special theory of relativity in the first thirty-eight sections. They will also be able to read Chapter VIII and in this way obtain an insight into the ideas underlying Einstein's general theory of relativity without spending any time on the laborious task of learning its special mathematical methods.

We have included only those developments of the theory of relativity which can be regarded as safely established, the various attempts at constructing a unified theory of gravitation and electromagnetism falling outside the scope of the present book. Also the cosmological problems have only been touched upon, since these problems have been extensively treated by Tolman in this series of monographs. Within these restrictions it is hoped, however, that the reader will find a fairly complete and well-rounded account of one of the most beautiful chapters in the history of science, which for the main part was written by a single man, Albert Einstein.

On completion of this work I gratefully acknowledge the help and advice which I have received from many quarters. First of all I want to express my deep gratitude to Professor Niels Bohr for his kind interest in my work during all these years and for the constant inspiration derived from many discussions and conversations at his institute. My thanks are due to Professor N. F. Mott and Professor I. N. Sneddon for reading the manuscript and eliminating the worst danicisms. I also wish to thank the staff of the Clarendon Press for their friendly co-operation.

I am indebted to Dr. W. Kohn and Dr. W. J. Swiatecki for many suggestions which have considerably improved the text and, in particular, to mag. scient. J. Lindhard who has been of great help in checking all the equations and reading the proofs. Finally, I am grateful to Miss S. Hellmann for her untiring assistance in the preparation of the manuscript and the proof-reading.

C. M.

COPENHAGEN
November 1951

CONTENTS

I

THE FOUNDATIONS OF THE SPECIAL THEORY OF RELATIVITY. HISTORICAL SURVEY

1. The relativity principle of mechanics. The Galilean transformation

THE special theory of relativity which was developed in the beginning of the twentieth century, especially through Einstein's work, has its roots far back in the past. In a way, this theory can be regarded as a continuation and completion of the ideas which have been the basis of our description of nature since the times of Galileo and Newton. The basic postulate of this theory, the so-called special principle of relativity,† had already in Galileo's and Huyghens's works played a decisive role in the development of the fundamental laws of mechanics. Also the validity of the principle of relativity for the phenomena of mechanics is a simple consequence of the Newtonian laws of mechanics. Since the laws of mechanics are especially well suited for the illustration of the principle of relativity, we shall start by considering purely mechanical phenomena.

According to Newton's first law, the law of inertia, a material particle when left to itself will continue to move in a straight line with constant velocity. Since one cannot simply speak of motion, but only of motion relative to something else, this statement has a precise meaning only when a certain well-defined system of reference has been established relative to which the velocity of the particle is assumed to be measured. Therefore Newton introduced the notion of the 'absolute space', representing that system of reference relative to which every motion should be measured. Experience shows that the fixed stars as a whole may be regarded as approximately at rest relative to the 'absolute space', for a body sufficiently far away from celestial matter always moves with uniform velocity relative to the fixed stars.

It is, however, obvious that the law of inertia holds also in every other rigid system of reference moving with uniform velocity relative to the 'absolute' system, for a free particle will also be in uniform translatory motion with respect to such a system. All systems of reference for which the law of inertia is valid are called systems of inertia. They form a

† When reference is made to the principle of relativity in Chapters I–VII we always have in view the principle of special relativity as contrasted with the principle of general relativity which is the basis of the general theory of relativity.

3595.60

threefold infinity of rigid systems of reference moving in straight lines and with constant velocity relative to each other. One of them is the absolute system which is at rest relative to the fixed stars as a whole; but as regards the validity of the law of inertia, all systems of inertia are completely equivalent.

Now, the principle of relativity in mechanics states that the systems of inertia are also completely equivalent with regard to the other laws of mechanics. If this is true all mechanical phenomena will take the same course of development in any system of inertia so that it is impossible from observations of such phenomena to detect a uniform motion of the system as a whole relative to the 'absolute' system. Thus, a study of mechanical phenomena alone can never lead to a determination of the 'absolute' system.

We shall now see that the fundamental equations of Newtonian mechanics actually are in accordance with the principle of relativity. Let us consider two arbitrary systems of inertia, I and I'. In each of these frames of reference we use definite systems of coordinates S and S'. We may, for instance, choose Cartesian coordinates $\mathbf{x} = (x, y, z)$ and $\mathbf{x}' = (x', y', z')$ in I and I', respectively. According to the conceptions of space and time, derived from our usual experience, which also form the basis of the Newtonian formulation of the fundamental laws of mechanics, the connexion between the coordinate vectors $\mathbf{x}$ and $\mathbf{x}'$ for one and the same space point in the two coordinate systems S and S' is given by

$$\mathbf{x}' = \mathbf{x} - \mathbf{v}t, \tag{1a}$$

where $\mathbf{v}$ is a vector denoting velocity and direction of motion of S' relative to S. t is the time and, for the sake of simplicity, it is assumed that the origins of the two systems of coordinates coincide at the time $t = 0$. To the equations (1 a) may be added the equation

$$t' = t \tag{1b}$$

which states that the parameter describing the time is the same in all systems of inertia. Thus, in the Newtonian description of physical phenomena the time is an absolute quantity. The equations (1 a) and (1 b) are often referred to as the *Galilean transformation.*

If the directions of the axes of the two systems of coordinates are parallel, and if $\mathbf{v}$ has the direction of the x-axis, we obtain a *special* Galilean transformation which can be written

$$\left.\begin{aligned} x' &= x - vt, \quad y' = y, \quad z' = z, \\ t' &= t \end{aligned}\right\}. \tag{2}$$

Since the systems of coordinates S and S' are completely equivalent, at any rate as far as kinematics is concerned, and since S obviously moves with the velocity $-\mathbf{v}$ relative to S', the inverse transformations to (1) and (2) are simply obtained by interchanging the primed and the unprimed variables and simultaneously replacing $\mathbf{v}$ by $-\mathbf{v}$.

Let us now consider an arbitrary motion of a material particle. By differentiation of (1 *a*) we get

$$\frac{d\mathbf{x}'}{dt'} = \frac{d\mathbf{x}}{dt} - \mathbf{v}$$

or

$$\mathbf{u}' = \mathbf{u} - \mathbf{v}, \tag{3}$$

where $\mathbf{u}$ and $\mathbf{u}'$ represent the velocities of the particle in the two systems of inertia. (3) is the usual addition theorem of velocities. For a special Galilean transformation (2), (3) reduces to

$$u'_x = u_x - v, \quad u'_y = u_y, \quad u'_z = u_z. \tag{4}$$

When the velocity vector $\mathbf{u}$, and thus also $\mathbf{u}'$, is perpendicular to the z-axis, (4) may be written

$$u' \cos\vartheta' = u \cos\vartheta - v,$$

$$u' \sin\vartheta' = u \sin\vartheta,$$

where ϑ and ϑ' are the angles between the x-axis and the directions of $\mathbf{u}$ and $\mathbf{u}'$, respectively. Further, $u = |\mathbf{u}|$, $u' = |\mathbf{u}'|$ denote the magnitudes of the vectors $\mathbf{u}$ and $\mathbf{u}'$. If we now divide one of these equations by the other we obtain

$$\tan\vartheta' = \frac{\sin\vartheta}{\cos\vartheta - v/u}, \tag{5}$$

and by summation of the squares of the equations we get

$$u' = u\left\{1 - 2\frac{v}{u}\cos\vartheta + \frac{v^2}{u^2}\right\}^{\frac{1}{2}}. \tag{6}$$

Now let us assume that the material particle with the mass m is acted on by a force $\mathbf{F}$. In the absolute system of coordinates S the particle will then obtain an acceleration given according to Newton's second law by the equation

$$m\frac{d^2\mathbf{x}}{dt^2} = \mathbf{F}. \tag{7}$$

From (1 *a*) and (1 *b*) it now follows that

$$\frac{d^2\mathbf{x}'}{dt'^2} = \frac{d^2\mathbf{x}}{dt^2}, \tag{8}$$

and since in Newtonian mechanics forces and masses are absolute quantities, i.e.

$$\mathbf{F}' = \mathbf{F}, \quad m' = m, \tag{9}$$

we obtain

$$m' \frac{d^2\mathbf{x}'}{dt'^2} = \mathbf{F}'. \tag{10}$$

Thus we see that the second law of Newton is valid in every system of inertia in accordance with the principle of relativity. This can be expressed more accurately by stating that the Newtonian fundamental equations are invariant under Galilean transformations. As is well known, this invariance does not hold for more general transformations leading to accelerated systems of reference. If one wants to treat mechanical phenomena in such systems, one has to introduce extra fictitious forces, e.g. centrifugal forces and Coriolis forces which only depend on the acceleration of the frame of reference and therefore are in no causal relationship with the physical properties of other terrestrial systems. It was just this difference between the uniformly moving and the accelerated systems of reference which led Newton to the conception of absolute space.

2. The special principle of relativity

As already mentioned, the validity of the principle of relativity in mechanics prevents a unique determination of the absolute system of reference from studies of mechanical phenomena alone. Now the basic assumption of the special theory of relativity is that *the special principle of relativity is valid for all physical laws.*† According to this theory, all physical phenomena should have the same course of development in all systems of inertia, and observers installed in different systems of inertia should thus as a result of their experiments arrive at the establishment of the same laws of nature.

If this is so, the notion of absolute space obviously loses its meaning, since any system of inertia with equally good reason can claim to be the absolute system of reference. Of course nobody can prevent us from calling one definite system of inertia, e.g. the one which is at rest relative to the fixed stars, the absolute system and expressing all laws of nature in coordinates of this system. Such a procedure is, however, extremely unsatisfactory in view of the arbitrariness in the choice of the absolute system. It is, furthermore, very inconvenient to proceed in this manner. The physical experiments from which the laws of nature are derived are usually not performed in a system of reference which is at rest relative

† With the exception of the laws of gravitation which find their natural place in the general theory of relativity.

to the fixed stars. On account of its motion around the sun the earth will in the course of a year represent widely different systems of inertia if we disregard the small acceleration of the earth in this motion. The transformation to the coordinates of the 'absolute' system is therefore rather complicated.

The validity of the principle of relativity for all physical phenomena now makes such a transformation unnecessary, since the system of inertia in which the earth is at rest at the moment considered is equivalent to any other system of inertia. This obviously leads to an enormous simplification in our description of nature.

However, this simplification has to be paid for, as we shall now see, by an abandonment of our usual notions of time and space. The extension of the principle of relativity to electromagnetic phenomena means, as mentioned before, that physicists who have established their laboratories in two different systems of inertia will, as a result of their experiments, be led independently to Maxwell's fundamental equations of electrodynamics. These equations contain a universal constant c which can be determined by means of purely electromagnetic measurements, and is very closely equal to 3×10^{10} cm./sec.† On the other hand, it is a simple consequence of Maxwell's equations that electromagnetic waves in empty space propagate with the velocity c, independently of the way in which they are created. Since light waves, according to Maxwell's theory of light, are special electromagnetic waves, the velocity with which light is propagated *in vacuo* must also be independent of the state of motion of the light source and equal to the constant c. If Maxwell's equations in accordance with the relativity principle are valid in any system of inertia, the velocity of light must have the same constant value c in all systems of inertia, independently of the motion of the light source. This is obviously in conflict with the usual kinematical concepts according to which we should expect, for instance, to find a lower velocity of light in S' than in S if the relative motion of S' with respect to S has the same direction as the direction of propagation of the light ray.

Consequently the acceptance of the relativity principle must necessarily lead to a revision of our ordinary concepts of space and time. Before taking such a radical step one would naturally want to be sure that it is really necessary. This question can only be settled as a result of experiment. Optical experiments are especially suited to this purpose in view of the high accuracy obtainable with optical instruments. In

† W. Weber und R. Kohlrausch (1856), *Ostwalds Klassiker der exakten Wissenschaften*, No. 142.

the following section we shall therefore give a short historical survey of the numerous optical experiments which have been performed in an attempt to detect effects depending on the motion of the apparatus with respect to an 'absolute' space. These experiments all gave negative results and finally led to a general acceptance of the principle of relativity.

3. Invariance of the phase of a plane wave

In contradistinction to the relativistic standpoint according to which Maxwell's equations are valid in any inertial system, Maxwell and his contemporaries maintained that the fundamental equations of electrodynamics were valid only in one system of inertia, viz the system which is at rest relative to the so-called 'world ether'. The ether was imagined as a medium which penetrates through all matter and empty space and which was the carrier of all optical and electromagnetic phenomena Moreover, the ether was supposed to represent the absolute system of reference, thus giving a substantial physical meaning to Newton's notion of absolute space. In the present section we shall fully adopt this point of view, and our first task will be to see what consequences this will have for the course of development of optical phenomena in a system of inertia moving relative to the ether.

Let S be a Cartesian coordinate system which is at rest in the ether. Relative to S, a plane monochromatic light wave in empty space will have the propagation velocity $c = 3\times 10^{10}$ cm /sec. A wave of this type is completely determined by the phase velocity, the frequency of the wave, and the direction of propagation. In the first place, we shall find the transformation of these three quantities by a transition to a coordinate system S' moving relative to the ether with a constant velocity v in the direction of the x-axis.

For simplicity let us assume that the normal of the wave plane lies in the xy-plane. Then the wave is described in S by a wave function

$$\psi = A\cos 2\pi F,$$
$$F = \nu\left(t-\frac{x\cos\alpha+y\sin\alpha}{c}\right) = \nu\left(t-\frac{l}{c}\right), \tag{11}$$

where ν is the frequency and α is the angle between the wave normal **n** and the x-axis. $l = x\cos\alpha+y\sin\alpha$ is the distance from the origin O to that wave plane which contains a point p with the coordinates (x, y) (cf. Fig. 1).

The phase F in (11) has the following simple physical meaning. Let us assume that the wave crest which passes the origin O at the time

$t = 0$ is provided with a label. Now suppose an observer to be placed at the point p who at the moment when the labelled wave arrives at p begins to count the waves passing over the point p. *The number of waves counted by the observer up to the time t will then just be equal to the phase F.* In fact, ν waves arrive per second and, since the labelled wave takes l/c seconds to move from O to p, the observer is counting during an interval of $t-(l/c)$ seconds.

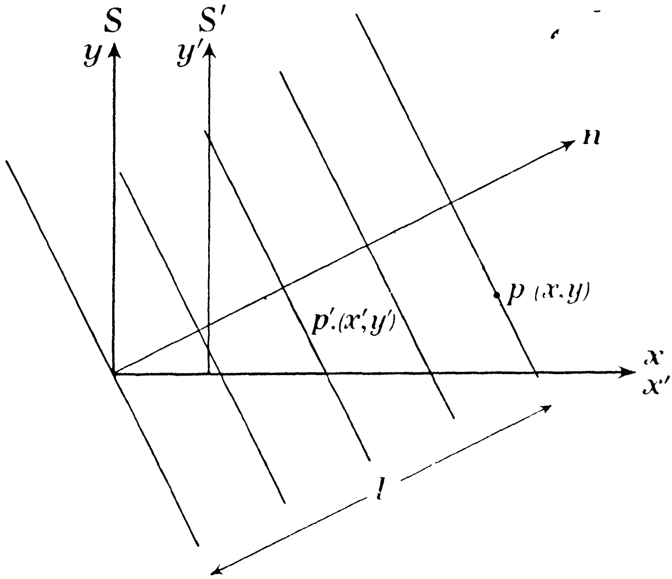

Fig. 1.

Now let S' be the moving system introduced above, and let us assume that the two coordinate systems S' and S coincide at the time $t = 0$ when the labelled wave passes the common origin of the systems. If p' is a point in S' with coordinates (x', y'), which coincides with p at the time $t = t'$, the number of waves passing p' from the time of arrival at p' of the labelled wave up to the time t will of course be the same number F as before. On the other hand, this number counted by an observer at p' will by a similar argument to that applied in S be equal to

$$F = \nu'\left(t'-\frac{l'}{c'}\right) = \nu'\left(t'-\frac{x'\cos\alpha'+y'\sin\alpha'}{c'}\right), \tag{12}$$

where the primed letters in (12) denote the same physical quantities as the corresponding unprimed letters in (11), but now measured in the system of coordinates S'. *Thus the phase F is an invariant.*

4. Transformation of the characteristics of a plane wave

The connexion between the coordinates (x, y, t) in (11) and the coordinates (x', y', t') in (12) is obviously given by the Galilean transformation (2), since the point p' coincides with p at the time $t' = t$. Equating the two expressions (11) and (12) for F and eliminating the coordinates (x, y, t) by means of (1) we obtain

$$\nu\left(t' - \frac{(x'+vt')\cos\alpha + y'\sin\alpha}{c}\right) = \nu'\left(t' - \frac{x'\cos\alpha' + y'\sin\alpha'}{c'}\right). \tag{13}$$

Now this equation must hold for all values of the independent variables x', y', t', and this is possible only when the coefficients of these variables are equal on both sides of the equation (13). Consequently we get the equations

$$\nu' = \nu\left(1 - \frac{v}{c}\cos\alpha\right), \tag{14}$$

$$\left.\begin{aligned} \frac{\nu'\sin\alpha'}{c'} &= \frac{\nu\sin\alpha}{c} \\ \frac{\nu'\cos\alpha'}{c'} &= \frac{\nu\cos\alpha}{c} \end{aligned}\right\}. \tag{15}$$

From equations (15) we get at once

$$\tan\alpha' = \tan\alpha, \tag{16}$$

i.e.

$$\alpha' = \alpha, \tag{17}$$

and, further,

$$\frac{\nu'^2}{c'^2} = \frac{\nu^2}{c^2}. \tag{18}$$

By solving the last equation for c' we find by means of (14) that

$$c' = c - v\cos\alpha. \tag{19}$$

The equations (14), (17), and (19) show how the three characteristics of the wave, viz. the frequency, the direction of the wave normal, and the phase velocity, will change at the transition to a coordinate system in uniform motion with respect to the ether. Equation (17) shows that the direction of the wave normal is the same in both systems of inertia. On the other hand, if $\alpha \neq \frac{1}{2}\pi$, the equations (14) and (19) involve the velocity v, so that a measurement of frequency and velocity in principle should be suited to determine the motion of the laboratory system with respect to the ether. In the following we shall discuss these two effects separately.

5. The Doppler effect

Equation (14), which is the mathematical expression of the so-called Doppler effect for light waves, gives the connexion between the frequency ν' in a moving frame of reference and the 'absolute' frequency ν as observed by an observer at rest in the ether. If $\mathbf{n}$ denotes a unit vector in the direction of the wave normal, and $\mathbf{v}$ is the velocity vector of S' relative to the ether, (14) can also be written

$$\nu' = \nu\left(1-\frac{\mathbf{n}.\mathbf{v}}{c}\right), \tag{20}$$

where $\mathbf{n}.\mathbf{v}$ is the scalar product of the two vectors. The Doppler effect appears when the observer is moving relative to the source of light. However, the formula (20) cannot be used directly in such a case, since as a rule both the observer and the source of light will have a motion relative to the ether. If ν^0 is the proper frequency of the light source, i.e. the frequency measured by an observer at rest relative to the source of light, we have in analogy to (20)

$$\nu^0 = \nu(1-\mathbf{n}.\mathbf{v}^0/c), \tag{21}$$

where $\mathbf{v}^0$ is the velocity of the source relative to the ether.

By elimination of the unknown absolute frequency ν we obtain from (20) and (21)

$$\nu' = \nu^0\,\frac{1-(\mathbf{n}.\mathbf{v})/c}{1-(\mathbf{n}.\mathbf{v}^0)/c}. \tag{22}$$

The frequencies ν' and ν^0, the direction of propagation of the light $\mathbf{n}$, and the relative velocity $\mathbf{v}_r = \mathbf{v}-\mathbf{v}^0$ of the observer with respect to the source of light can be determined directly by experiments, and (22) then permits in principle a determination of the absolute velocities $\mathbf{v}^0$ and $\mathbf{v}$ of the light source and the observer.

Both $\mathbf{v}$ and $\mathbf{v}^0$ are, however, very small compared with c, so that we may perform an expansion of (22) in terms of the small quantity $(\mathbf{v}^0\mathbf{n})/c$. If we neglect all terms of higher than second order in v/c and v^0/c, we get, introducing the relative velocity $\mathbf{v}_r = \mathbf{v}-\mathbf{v}^0$,

$$\nu' = \nu^0\left\{1-\frac{(\mathbf{n}.\mathbf{v}_r)}{c}-\frac{(\mathbf{n}.\mathbf{v}^0)(\mathbf{n}.\mathbf{v}_r)}{c^2}\right\}. \tag{23}$$

In first approximation, the Doppler effect thus depends only on the relative velocity $\mathbf{v}_r$. The absolute velocity $\mathbf{v}^0$ of the light source enters only in the small second-order terms.

The Doppler effect is observed in the spectra of the stars, the lines of the spectrum being shifted towards violet or red according as the earth, during its annual motion, moves nearer to or away from the observed star. The velocity of the earth in its orbit is approximately 3×10^6 cm./sec.

and cosmic velocities are mostly of the same order of magnitude. Consequently, we have $v/c \approx 10^{-4}$, i.e. the terms of the second order will be of the order of magnitude 10^{-8}, which is far beyond the precision of such measurements.

The Doppler effect has been observed also in the light from moving terrestrial sources. By measuring the frequency of the light emitted by rapidly moving hydrogen molecule ions in a positive ray tube, J. Stark† found good agreement with the formula (23) as regards the terms of the first order. In these experiments the relative velocity v_r and, thus, also v^0 were of the order of 10^8 cm./sec., i e. $v/c \approx \frac{1}{300}$. Also in this case, the second-order terms were too small to be measured, and these experiments did not therefore allow a determination of the absolute velocity.

Much later, in the thirties, such experiments were repeated by Ives‡ with an improved experimental technique, which allowed the second-order terms to be determined. The results obtained were *not* in agreement with (23), but they agreed with a formula derived from the theory of relativity (equation (II. 90), Chap. II, § 25). The second-order term was found to be independent of the direction of the light emitted and dependent only on the *relative* velocity v_r. No motion relative to the ether could be observed, in agreement with the principle of relativity.

These experiments, which were performed much later, had of course no influence on the historical development of the relativity theory. Because of the limited accuracy, the experiments by Stark did not allow any decision to be formed on the validity of the principle of relativity, but on the other hand the results of these experiments were not in contradiction with the principle of relativity.

6. The velocity of light *in vacuo*

We now turn to the question whether a measurement of the velocity of light by terrestrial methods can lead to a determination of the absolute velocity v of the earth. Since v enters into equation (19), which can also be written

$$c' = c-(\mathbf{n}.\mathbf{v}), \tag{24}$$

this should be possible in principle, as mentioned on p. 8. The well-known measurements of the velocity of light by Fizeau (1849) and Foucault (1865) showed, however, no influence at all of the motion of the earth. The velocity of light was always found to be the same, in agreement

† J. Stark, *Ann. d Phys* **21**, 40 (1906), J. Stark and K Siegel, ibid **21**, 457 (1906), J Stark, W. Hermann, and S. Kinoshita, ibid **21**, 462 (1906).

‡ Cf , for example, H. E Ives and G. R Stilwell, *Journal of the Optical Society of America*, **28**, 215 (1938).

with the principle of relativity. How can this result be understood on the basis of the ether theory?

In the first place, it should be observed that what is measured in these experiments is not the phase velocity but the so-called ray velocity. In Fizeau's original experiments, for example, a light signal is sent along a certain path and back again, and the difference between the time of departure and the time of return of the signal is measured. The velocity is then determined as the ratio between the length of the path traversed and this time interval. Now it is true that the velocity of the light signal is equal to the phase velocity c in the coordinate system S which is at rest in the ether, but in the moving system S' the velocity of the signal will not be equal to the phase velocity c' given by (24). This is plausible if one keeps in mind that a light signal represents a certain amount of electromagnetic energy, and that energy, like mass, is a quantity which is conserved, so that a signal in many respects will resemble a material particle. Therefore we should rather expect that the velocity of a light signal in a moving system of coordinates S' is given by the equations (3), (5), and (6) if in these equations u is put equal to c, the velocity of light in the ether.

A closer treatment based on the wave theory of light confirms this expectation. Such a treatment shows that in a moving coordinate system the ether acts like an anisotropic medium so that one has to distinguish between the phase velocity and the ray velocity, which is identical with the velocity of propagation of the light energy and is just given by (3). To see this we shall apply the well-known Huyghens principle which for all phenomena in the domain of geometrical optics is a consequence of Maxwell's equations. In accordance with this principle we obtain the consecutive wave surfaces as the envelopes of elementary waves starting from each point of a wave surface.

Let us consider the propagation of light in a system of coordinates S' moving with the velocity $\mathbf{v}$ relative to the ether system S. In S' we thus have an 'ether wind' of the velocity $-\mathbf{v}$ which will carry along the elementary waves in the same way as sound waves are carried by the wind.

Fig. 2 gives a diagram of successive positions of light waves in the system S'. Let the surface σ denote the position of a given wave front at the time t. In order to construct the wave front σ_1 at the time $t+dt$ we regard every point P on σ as a starting-point of an elementary wave. Because of the ether wind, this elementary wave will at the time $t+dt$ obviously form a sphere E with centre at a point Q which lies at a distance

$v\,dt$ from the point P in the direction of the ether wind. The infinitesimal vector $\overrightarrow{PQ}$ is thus given by

$$\overrightarrow{PQ} = -\mathbf{v}\,dt. \tag{25}$$

Since the velocity of propagation of the elementary wave in the ether is c, the sphere E has a radius $QP_1 = c\,dt$. The wave front σ_1 is now

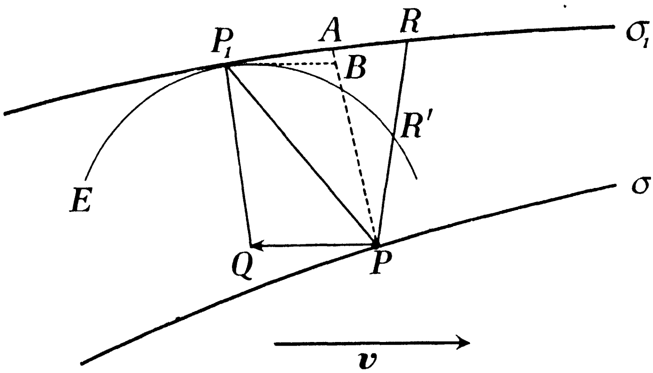

FIG 2.

obtained as the envelope of all the elementary waves, i.e. the vector $\overrightarrow{QP_1}$ is perpendicular to σ_1 at the point of contact P_1, and in the limit as $dt \to 0$ $\overrightarrow{QP_1}$ is also perpendicular to σ. Thus, the vector $\overrightarrow{QP_1}$ lies in the direction of the phase-velocity vector, and we have consequently

$$\overrightarrow{QP_1} = \mathbf{c}\,dt = c\mathbf{n}\,dt, \tag{26}$$

where $\mathbf{c} = \mathbf{n}c$ is the phase-velocity vector in the ether. In S' the phase velocity is by definition given by

$$\overrightarrow{PA} = \mathbf{c}'\,dt = c'\mathbf{n}'\,dt, \tag{27}$$

where the unit vector $\mathbf{n}'$ denotes the direction of the wave normal in S'. Obviously

$$\mathbf{n}' = \mathbf{n} \tag{28}$$

in accordance with (17). Since any infinitesimal part of a curved wave surface can be regarded as plane, the connexion between c' and c must again be given by (24). This follows also directly from (27), (28), (26), and (25) if we note that

$$PB = QP_1 = c\,dt,$$

while BA is equal to the projection of the vector $\overrightarrow{BP_1} = \overrightarrow{PQ} = -\mathbf{v}\,dt$ on the direction $\mathbf{n}$.

The relative direction of the ray, i.e. the direction of propagation of the light energy as estimated by an observer in S', is now given by the direction of the vector $\overrightarrow{PP_1}$, and if $\mathbf{u}'$ denotes the relative velocity of the ray we have

$$\overrightarrow{PP_1} = \mathbf{u}'\,dt = u'\mathbf{e}'\,dt, \tag{29}$$

where $\mathbf{e}'$ is a unit vector indicating the relative direction of the ray.

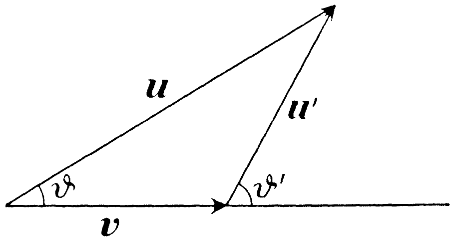

FIG. 3

Now the vector $\overrightarrow{P'P_1}$ is the sum of the vectors $\overrightarrow{PQ}$ and $\overrightarrow{QP_1}$

$$\overrightarrow{PP_1} = \overrightarrow{PQ} + \overrightarrow{QP_1}. \tag{30}$$

In the limit as $dt \to 0$ this gives, because of (25), (26), and (29),

$$\mathbf{u}' = \mathbf{c} - \mathbf{v}. \tag{31}$$

In the absolute system the ray velocity is identical with the phase velocity

$$\mathbf{u} = u\mathbf{e} = \mathbf{c} = c\mathbf{n}, \tag{32}$$

i.e. $$u = c, \qquad \mathbf{e} = \mathbf{n} = \mathbf{n}',$$

so that (31) can be written

$$\mathbf{u}' = \mathbf{u} - \mathbf{v}. \tag{33}$$

Thus we obtain the same addition theorem for ray velocities as for particle velocities· $\mathbf{u}$ is the geometrical sum of $\mathbf{u}'$ and $\mathbf{v}$. If ϑ and ϑ' are the angles between the direction of the velocity vector $\mathbf{v}$ and the absolute and relative ray directions, respectively, a consideration of the triangle in Fig. 3 gives at once, since $u = c$,

$$\tan\vartheta' = \frac{\sin\vartheta}{\cos\vartheta - v/c}, \tag{34}$$

and $$u'^2 + v^2 + 2vu'\cos\vartheta' = u^2 = c^2.$$

By solving this equation with respect to u' we get

$$\begin{aligned} u' &= \{c^2 - v^2 + v^2\cos^2\vartheta'\}^{\frac{1}{2}} - v\cos\vartheta' \\ &= \{c^2 - v^2 + (\mathbf{v}.\mathbf{e}')^2\}^{\frac{1}{2}} - (\mathbf{v}.\mathbf{e}'). \end{aligned} \tag{35}$$

A comparison between (35) and (24) shows that the relative ray velocity in general is different from the relative phase velocity, the difference being of the second order in v/c. Only when the direction of the ray (and the direction of the wave normal) is equal or opposite to the direction of **v**, are the two velocities identical and equal to $c-v$ and $c+v$, respectively.

Now it is clear that the velocity measured by Fizeau's and Foucault's methods is the ray velocity, but, since **v** also enters into (35), it should be possible in principle to determine the absolute velocity of the earth on the basis of these measurements. It is, however, easy to understand why no variation in the velocity of light was ever observed. In these experiments a light ray is sent along a known closed path and the time which the light signal takes to travel along this path is measured. In order to make the path as long as possible within a limited space, the light ray is reflected many times by suitably arranged mirrors. Let $l_1, l_2,..., l_\iota$ be the distances traversed by the ray between the mirrors, and let the corresponding directions of the ray be given by the unit vectors $\mathbf{e}'_1, \mathbf{e}'_2,..., \mathbf{e}'_\iota$; then we obviously have

$$\sum_i l_\iota \mathbf{e}'_\iota = 0, \tag{36}$$

since the ray describes a closed polygon. The time needed for the light to traverse this closed path is then, according to (35),

$$t = \sum_\iota \frac{l_\iota}{\{c^2-v^2+(\mathbf{v}\cdot\mathbf{e}'_\iota)^2\}^{\frac{1}{2}}-(\mathbf{v}\cdot\mathbf{e}'_\iota)}. \tag{37}$$

If this expression is expanded in terms of the small quantity v/c we get, neglecting terms of order higher than the first,

$$t = \sum_\iota \frac{l_\iota}{c} + \sum_\iota \frac{l_\iota(\mathbf{v}\cdot\mathbf{e}'_\iota)}{c^2}.$$

On account of (36), the first-order term disappears and in this approximation we obtain

$$t = \sum_\iota l_\iota/c.$$

Thus, when terms of order higher than the first are neglected, the measured time is the same as if the earth were at rest in the ether. A determination of the absolute velocity **v** would consequently require a measurement of quantities of at least second order. Fizeau's and Foucault's methods, however, did not allow such a high accuracy, and it is therefore understandable even on the basis of the ether theory that the experimental results were in agreement with the principle of relativity. It was not until many years later that Michelson was able to develop a method which also allowed the measurement of magnitudes of the second order and thereby gave a final proof of the validity of the principle of relativity. We shall, however, follow the historical trend of development and return to a discussion of Michelson's experiments in a later section (§ 12).

7. The velocity of light in refractive media

Hitherto we have only discussed the propagation of light in empty space. Let us now assume that the space is filled with an isotropic transparent substance with the index of refraction n. If the substance is at rest relative to the world ether, the phase velocity in the absolute system S is, according to Maxwell's phenomenological electrodynamics,

$$c_1 = \frac{c}{n}, \qquad n = (\epsilon\mu)^{\frac{1}{2}}, \tag{38}$$

where ϵ is the dielectric constant of the medium and μ its magnetic permeability The phase velocity c_1' relative to a moving system of co-ordinates is then, in analogy to (24), given by

$$c_1' = c_1 - (\mathbf{n}.\mathbf{v}). \tag{39}$$

This formula is valid if the refractive body is at rest in the absolute system S But suppose the body moves with a velocity $\mathbf{v}$, thus being at rest in S', what will then be the expression for the phase velocity in S'?

This problem was much discussed in the early days. The simplest assumption is that (39) remains valid, i.e that the ether passes undisturbed through the moving body without being dragged along. Then we have in S' an ether wind with the velocity $-\mathbf{v}$, and we can now find the ray velocity by means of Huyghens's principle in the same way as in § 6. In equation (35) we have simply to replace c by the phase velocity $c_1 = c/n$ and thus get the following expression for the ray velocity in S':

$$u' = \{c_1^2 - v^2 + (\mathbf{v}.\mathbf{e}')^2\}^{\frac{1}{2}} - (\mathbf{v}.\mathbf{e}'). \tag{40}$$

Instead of assuming that the ether passes undisturbed through the moving body, it has also been suggested that the ether is completely dragged along by the body. According to this hypothesis, which was put forward by Stokes,† we obviously get

$$u' = c_1' = \frac{c}{n}, \tag{41}$$

for, in this case, there would be no ether wind in S'.

† G. G Stokes, *Phil Mag* (3), **27**, 9 (1845), *Mathematical and Physical Papers*, **1**, 134 (1880).

A third possibility is to assume that the ether is dragged along only partly by the moving body, say with a velocity $\alpha\mathbf{v}$, where the 'dragging coefficient' α is a positive number smaller than 1 depending on the refractive index n. This hypothesis was put forward by Fresnel† who, on the basis of the elastic ether theory, gave the following expression for the dragging coefficient:

$$\alpha = 1 - \frac{1}{n^2}. \tag{42}$$

On this hypothesis the velocity of S' relative to the dragged ether is $\mathbf{v} - \alpha\mathbf{v} = (\mathbf{v}/n^2)$ and, instead of (39), we obtain for the relative phase velocity

$$c_1' = c_1^* - \frac{(\mathbf{n}.\mathbf{v})}{n^2}, \tag{43}$$

where $c_1^* = c/n$ is the phase velocity in a system of coordinates S^* accompanying the dragged ether. Consequently the phase velocity in the absolute system S is

$$c_1 = c_1^* + \alpha(\mathbf{v}.\mathbf{n}) = \frac{c}{n} + (\mathbf{v}.\mathbf{n})\left(1 - \frac{1}{n^2}\right), \tag{44}$$

since S is moving with the velocity $-\alpha\mathbf{v}$ relative to S^*.

In order to find the relative ray velocity u' by means of the method outlined in § 6, we must let the system of coordinates S^* take over the role played by S in the former considerations. Since the ether wind in S' has the velocity $-(\mathbf{v}/n^2)$, we get, in analogy to (31),

$$\mathbf{u}' = \mathbf{c}_1^* - \frac{\mathbf{v}}{n^2}, \tag{45}$$

where the vector $\mathbf{c}_1^*$ with the magnitude $c_1^* = c/n$ is the phase-velocity vector in S^*. The magnitude of the relative ray velocity we obtain in the same way from (35) by replacing c and $-\mathbf{v}$ by $c_1^* = c/n$ and $-(\mathbf{v}/n^2)$, respectively, i.e.

$$u' = \{c_1^{*2} - v^2/n^4 + (\mathbf{v}.\mathbf{e}')^2/n^4\}^{\frac{1}{2}} - (\mathbf{v}.\mathbf{e}')/n^2. \tag{46}$$

In the absolute system S, however, we have an ether wind with the velocity $\alpha\mathbf{v} = (1 - 1/n^2)\mathbf{v}$. Therefore we have for the absolute ray velocity $\mathbf{u}$, in analogy to (45) and (46),

$$\left.\begin{aligned} \mathbf{u} &= \mathbf{c}_1^* + \alpha\mathbf{v}, \quad u = \{c_1^{*2} - \alpha^2 v^2 + \alpha^2(\mathbf{v}.\mathbf{e})^2\}^{\frac{1}{2}} + \alpha(\mathbf{v}.\mathbf{e}) \\ \alpha &= 1 - \frac{1}{n^2}, \quad c_1^* = \frac{c}{n}, \end{aligned}\right\} \tag{47}$$

† A. J. Fresnel, *Ann. de chim. et de phys.* **9**, 57 (1818).

or, neglecting terms of the second order in v,

$$u = \frac{c}{n} + \alpha(\mathbf{v}.\mathbf{e}). \tag{48}$$

By elimination of $\mathbf{c}_1^*$ from (45) and (47) we obtain again the simple addition theorem

$$\mathbf{u} = \mathbf{u}' + \mathbf{v}. \tag{49}$$

A comparison between the equations (43), (46), (44), and (47) shows that the ray velocities are identical with the phase velocities if the direction of propagation of the light is the same as or opposite to the direction of the velocity $\mathbf{v}$. Without any calculation this follows immediately from the fact that, in this case, the ether wind carries the elementary waves in a direction parallel to the light beam. The considerations given in this section are valid also for the case of inhomogeneous bodies with a continuously varying index of refraction. Only in this case the system S^*, which depends on the value of n, will be different at different points of the substance.

8. Hoek's and Fizeau's experiments

A measurement of the velocity of light in transparent substances seems to offer a new possibility for a determination of the absolute motion of the earth. An experiment of this kind was performed in 1868 by Hoek† who used an interferometer arrangement of the type shown in Fig. 4. A monochromatic light ray from a source of light L is divided by a (weakly silver-coated) glass plate P which is placed at an angle of 45° into a transmitted part 1 and a reflected part 2. The transmitted ray 1 is reflected by the mirrors S_1, S_2, S_3 and traverses a rectangular path $PS_1S_2S_3P$; again a certain fraction of the ray passes the plate P and enters the telescope T. The reflected ray 2 traverses the same rectangle in the opposite direction. On its return to P it is partly reflected into T where it interferes with 1. Between S_2 and S_3 is inserted a tube of length l filled with a substance with refractive index n (for instance water).

Even if the whole apparatus were at rest in the ether, such an arrangement would give rise to interference fringes in the telescope, since the slope of the mirrors cannot possibly be adjusted so accurately that two rays 1 and 2 which focus on the same point in the telescope have traversed a path of exactly the same optical length. However, if the whole apparatus has a velocity $\mathbf{v}$ with respect to the ether, this will cause an extra phase difference ΔF between the rays 1 and 2, which can be calculated by means of (40), (41), or (46).

† M. Hoek, *Archives Néerlandaises des Sciences Exactes et Naturelles*, **3**, 180 (1868).

Let us, for simplicity, assume that the apparatus is set up in such a way that the lines PS_1 and S_2S_3 are parallel to the direction of motion $\mathbf{v}$ of the apparatus relative to the world ether. The phase difference ΔF resulting from the absolute motion of the apparatus will then obviously be due to a difference in the times t_1 and t_2 which the rays 1 and 2, respectively, require to traverse the tube of length AB and the corresponding distance CD on the path connecting P and S_1. On the remaining paths DS_1S_2B and AS_3PC the two rays are completely equivalent so that no contribution to the phase difference ΔF can arise from these parts.

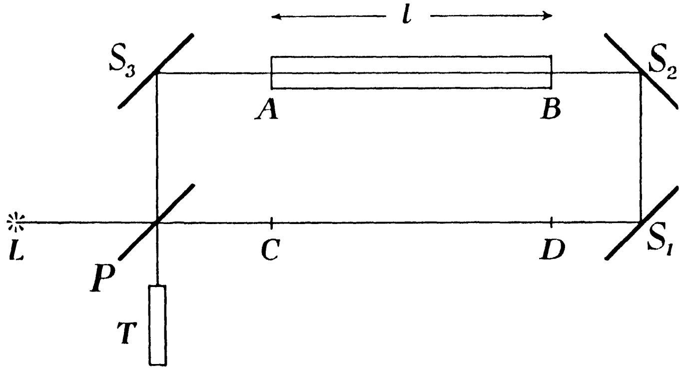

FIG 4.

The values of the quantities t_1 and t_2 will now depend on the degree to which the ether is dragged along by the refractive substance. If there is no dragging at all we get, according to (40),

$$t_1 = \frac{l}{c-v}+\frac{l}{c/n+v}, \qquad t_2 = \frac{l}{c/n-v}+\frac{l}{c+v},$$

where we have put the refractive index of air equal to 1. The phase difference between the rays 2 and 1, in so far as it is due to the absolute motion of the apparatus, is then

$$\Delta F = \nu(t_2-t_1) = 2l\nu v\left(\frac{1}{(c/n)^2-v^2}-\frac{1}{c^2-v^2}\right),$$

or, if we neglect terms of order higher than the first in v/c,

$$\Delta F = \frac{2l\nu}{c}\frac{v}{c}(n^2-1). \tag{50}$$

When the apparatus is at rest with respect to the earth, $\mathbf{v}$ is identical with the absolute motion of the earth. If, therefore, the apparatus were rotated 180° around an axis perpendicular to the direction of motion of the earth, the phase difference in question would be $-\Delta F$. Such a rotation should thus cause a shift of the interference lines corresponding to a phase shift of $2\Delta F$ and, since ΔF according to (50) is a quantity of the first order in v/c, this effect should be easily observable.

The result of Hoek's experiments was, however, negative; no observable shift of the interference lines could be detected after a rotation of the apparatus. This result too is in complete agreement with the principle of relativity according to which all phenomena should be independent of the state of motion of the measuring instruments.

Since the accuracy of Hoek's experiments did not go beyond terms of the first order, the negative result of this experiment was, however, not a serious difficulty for the ether theory; it only showed that the equation (40), based on the assumption that the ether is not dragged along by the refractive body, could not be maintained. Similarly, Stokes's hypothesis is ruled out by the result of Hoek's experiment, for from (41) we would get for ΔF

$$\Delta F = -\frac{2l\nu}{c}\frac{v}{c}. \tag{51}$$

On the other hand, the formulae (43) and (45), corresponding to a dragging coefficient (42), are seen to give an explanation of the result obtained by Hoek. In this case, we get, by means of (45) or (46),

$$\begin{aligned}\Delta F &= \nu\left(\frac{l}{c/n-v/n^2}+\frac{l}{c+v}-\frac{l}{c/n+v/n^2}-\frac{l}{c-v}\right)\\ &= 2l\nu v\left(\frac{1}{c^2-(v/n)^2}-\frac{1}{c^2-v^2}\right),\end{aligned} \tag{52}$$

and if we neglect terms of order higher than the first, this quantity is zero. It is also seen immediately that only Fresnel's value (42) for the dragging coefficient α gives a zero value for ΔF in first approximation. Hoek's experiment can therefore be regarded as an experimental verification of Fresnel's formulae for the velocity of light in moving bodies, at least as regards terms of first order.

As early as 1851, Fizeau† had obtained the same result by measuring the velocity of light in running water. The experimental arrangement was very similar to the interferometer arrangement in Hoek's experiment (see Fig. 5). The only difference is that the light rays 1 and 2 here are passing through water on the path PS_1 as well as on the path S_2S_3. As indicated in the figure, the ray 1 is traversing the water in a direction opposite to the direction of motion of the water, while the ray 2 has the same direction of motion as the water. Now Fizeau compared the position of the interference fringes while the water was at rest in the tubes

† H. Fizeau, *C.R.* **33**, 349 (1851); A. A. Michelson and E. W. Morley, *Amer. Journ. of Science*, **31**, 377 (1886); H. Fizeau, *Ann. d. Phys. und Chem., Erg.* **3**, 457 (1853).

with that when a strong water current was sent through the tubes. A marked shift of the fringes could be observed.

From Hoek's experiment we know that the motion of the earth relative to the ether can have an effect of second order only; we may therefore in our calculation make the assumption that the apparatus is at rest in the absolute system S. The velocity vector **v** in (48) is then simply equal

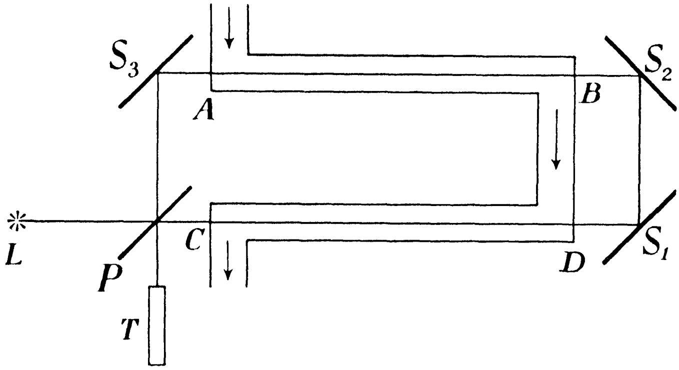

FIG. 5.

to the velocity of the water relative to the tube. Since **e** is parallel to **v** on the paths AB and CD, which are the only paths which give rise to a phase difference between 1 and 2, we obtain from (48) for this phase difference

$$\left.\begin{aligned} \Delta F &= l\nu[\{c/n-v(1-1/n^2)\}^{-1}-\{c/n+v(1-1/n^2)\}^{-1}] \\ &= \frac{2l\nu v}{c^2}(n^2-1), \end{aligned}\right\} \tag{53}$$

where l means the whole path through which the light ray travels in the water. The shift in the position of the fringes observed by Fizeau was in complete agreement with the phase difference given by (53).

9. Lorentz's theory of electrons

The experiments mentioned in the preceding section may be regarded as a decisive experimental verification of Fresnel's formulae (42)–(48), at least as regards all terms of first order. However, the derivation of these formulae from the point of view of the primitive ether theory meets with a serious difficulty when we keep in mind that the index of refraction generally depends on the frequency of the observed light. Since the dragging coefficient α in (42) is a function of n, this would mean that the dragging of the ether is not only dependent on the properties of the moving body, but also on the frequency of the light. Strictly speaking, one would have to introduce a separate ether for each colour of the light.

This is, of course, an impossible assumption, and this difficulty underlies both Fresnel's mechanical ether theory and Maxwell's phenomenological theory in which the ether represents the system of reference where Maxwell's equations are valid. In fact, the dependence of the dragging phenomena on the frequency makes it impossible to decide in which system of reference the equations of Maxwell's electrodynamics are valid.

This difficulty is closely connected with the fact that the index of refraction in this theory is constant and equal to $(\epsilon\mu)^{\frac{1}{2}}$ and thus does not give any explanation of the dispersion phenomena. A satisfactory explanation of the dispersion and of the dragging phenomena was given by Lorentz† in his theory of electrons. According to this theory, the ether is not dragged at all by refractive substances, but stays constantly at rest in a certain system of inertia—the absolute system. The material bodies are assumed to be composed of atoms which contain a number of positively and negatively charged electric particles. While the positive particles contain practically the whole mass of the atom, the negative particles, the electrons, are supposed to be very light. Under the influence of the electromagnetic fields in a light wave they perform forced vibrations around their equilibrium positions. Therefore the electrons themselves will emit electromagnetic waves which interfere with the incident wave in such a way that the effective velocity of propagation of the light in a medium at rest is c/n instead of c.

According to this theory it is also clear that the coefficient n in general depends on the position of the frequency of the incident wave relative to the proper frequencies of the electrons. Furthermore, Lorentz was able to show that a uniform motion of the refractive body modifies the waves emitted by the vibrating electrons in such a way that the effective phase velocity of the light in the moving body to a first approximation is given by Fresnel's formulae (43) and (44).

Thus as regards the propagation of light in refractive bodies Lorentz's electron theory gave, at least to a first approximation, the same results as Fresnel's theory, avoiding, however, the serious objections which could be raised against Fresnel's derivation of his formula. In one respect it even gave a more precise formulation of Fresnel's formula (44). Since the index of refraction n in dispersive media depends on the frequency ν, and since the frequencies on account of the Doppler effect are different

† H. A. Lorentz, *The Theory of Electrons*, Leipzig, 1916. See also L. Rosenfeld, *Theory of Electrons*, Amsterdam, 1951.

in S and in S', it must be specified which value for the frequency and thus for n should be inserted in (44). Now Lorentz was able to show that one must insert the value $n(\nu')$, where ν' is the frequency in the system S' moving along with the refractive body, while n is the function of the frequency which is valid for a body at rest in the ether.

We shall not go deeper into Lorentz's theory since all the results mentioned above will be derived later in a much simpler way from the theory of relativity (Chapter II). Here we shall confine ourselves to the remark that Fresnel's formulae (42)–(48) are a consequence of the electron theory if we neglect all terms of order higher than the first. In the following section we shall use these formulae to show that, for *all* optical effects of the first order, the ether theory in the form given it by Lorentz yields results which are in agreement with the postulate of relativity.

10. Agreement between the ether theory and the principle of relativity as regards all effects of the first order. Fermat's principle

According to the principle of relativity the track of a light ray connecting two points which are fixed relative to the earth should be completely independent of the absolute motion of the earth. This must at least be true approximately, otherwise it would be impossible to make constant optical images of objects. If the passage of the light rays through the lens systems of optical instruments were markedly dependent on the absolute motion of the earth, the image formation in such an instrument would be time-dependent, an effect which has, however, never been observed.

As was shown by Lorentz,† this fact can easily be explained on the basis of Lorentz's electron theory when we assume that all terms of the second order are too small to be measured. This result of Lorentz's theory is the more remarkable as the relative ray velocity u' is, to a first approximation, essentially dependent on the absolute velocity $\mathbf{v}$. Neglecting all terms of higher than the first order in (46), we get

$$u' = c/n-(\mathbf{v}\,.\,\mathbf{e}')/n^2. \tag{54}$$

In order to construct the track of a ray on the basis of this expression we shall again consider Fig. 2 (p. 12). As in (29), we have

$$\overrightarrow{PP_1} = u'\mathbf{e}'\,dt,$$

where u' is given by (54) and $\mathbf{e}'$ is a unit vector in the direction of the relative ray through the point P. But instead of (25) and (26) we have

† See ref., p. 21.

now $QP_1 = c\, dt/n$ and $\overrightarrow{PQ} = -\mathbf{v}\, dt/n^2$, where n is the index of refraction at the place in the medium considered, for in the present consideration the system S^*, which is at rest in the 'dragged' ether, plays the same role as the absolute system S *in vacuo*. Let us call two points P and P_1, which are lying on the same ray and on consecutive wave planes, σ and σ_1, conjugated points. Then the construction of the light ray obviously consists in a determination of the conjugated points on the consecutive wave planes. Now consider two arbitrary points on σ and σ_1, respectively, with the distance ds and let us form the quantity ds/u', where u' is given by (54), with the direction $\mathbf{e}'$ equal to the direction of the line connecting these points. If the two points are conjugated as P and P' in Fig. 2, ds/u' is equal to dt, where dt is the time which the wave front takes to travel from the position σ to the position σ_1. If the two points are not conjugated as, for instance, P and R in Fig. 2, ds/u' will always be larger than dt. For, in this case, we have $dt = PR'/u'$, where PR' is the distance from P to the point of intersection between the line PR and the elementary wave E, and since R lies outside E, we have $ds > u'\, dt$.

Now let A and B be two fixed points in the refractive body. Consider the integral

$$\int_A^B \frac{ds}{u'} \tag{55}$$

along an arbitrary curve connecting the points A and B, where u' at any point of the curve is the relative ray velocity (54) corresponding to the direction of the element ds. According to the above arguments the integral (55) will then assume the lowest value when the curve coincides with the light ray through the points A and B, because only in this case will all elements ds of the curve connect conjugated points.

The ray between two arbitrary fixed points A and B in the refractive body is thus determined by the condition that *the integral* (55) *is a minimum* for the track of the ray, and since the integral is equal to the time which the light ray needs to travel from A to B, this theorem is identical with *Fermat's principle* which is thus a consequence of Huyghens's principle.

For the integrand in (55) we get, by means of (54), as a first approximation

$$\frac{1}{u'} = \frac{1}{c/n-(\mathbf{v}\,.\,\mathbf{e}')/n^2} = \frac{n}{c}+\frac{(\mathbf{v}\,.\,\mathbf{e}')}{c^2}. \tag{56}$$

Hence,

$$\int_A^B \frac{ds}{u'} = \frac{1}{c}\int_A^B n\, ds+\frac{1}{c^2}\left(\mathbf{v}.\int_A^B \mathbf{ds}\right), \tag{57}$$

where $\mathbf{ds} = \mathbf{e}' \, ds$ is an infinitesimal vector joining two consecutive points on the curve connecting A and B. The last term in (57) is consequently equal to v/c^2 times the projection of the curve on the direction of $\mathbf{v}$ and this projection is the same for all curves connecting the fixed points A and B.

Thus, for the purpose of variation we can replace (55) by

$$\int_A^B \frac{ds}{u'} = \int_A^B \frac{n}{c} \, ds. \tag{58}$$

This expression is, however, equal to the time which the ray would need to travel from A to B if the refracting body were at rest in the ether. To a first approximation the light track between two points in the moving body is thus the same as if the body were at rest, in agreement with the principle of relativity. If the path of the ray in a medium at rest is mapped out by means of suitably arranged screens with small openings, the ray will also pass through these openings if the whole apparatus is moving with constant velocity.

Hoek's experiment showed that the interference phenomenon occurring in his special experimental arrangement, at least in first approximation, was independent of the absolute motion of the earth. Nor has any influence of the absolute motion of the earth ever been detected in the numerous later interference experiments. These facts, which are in complete agreement with the postulate of relativity, can, however, as shown by Lorentz,† easily be explained on the basis of the ether theory if we may assume that terms of the second order are below the accuracy of the experiments.

Let us consider an arbitrary interference experiment where all parts of the apparatus are at rest relative to the system of reference S' which follows the motion of the earth. Such an experiment always involves two rays 1 and 2 which start from the same point A and are brought to interference at another point B after having traversed different paths I and II from A to B. Now, according to (57), the times t_1 and t_2 which the rays 1 and 2 take to travel from A to B are

$$\left.\begin{aligned} t_1 &= \int_{\mathrm{I}} \frac{ds}{u'} = \int_{\mathrm{I}} \frac{n}{c} \, ds + \frac{1}{c^2}\left(\mathbf{v} \cdot \int_{\mathrm{I}} \mathbf{ds}\right) \\ t_2 &= \int_{\mathrm{II}} \frac{ds}{u'} = \int_{\mathrm{II}} \frac{n}{c} \, ds + \frac{1}{c^2}\left(\mathbf{v} \cdot \int_{\mathrm{II}} \mathbf{ds}\right) \end{aligned}\right\}, \tag{59}$$

† See ref., p. 21.

where the integrals occurring in (59) should be taken along the two paths I and II from A to B. Since these paths have common end-points, the last terms in the equations (59) are equal and the difference in time is simply given by

$$\Delta t = t_1 - t_2 = \int\limits_{\mathrm{I}} \frac{n}{c}\,ds - \int\limits_{\mathrm{II}} \frac{n}{c}\,ds. \tag{60}$$

Thus the time difference is the same as if the apparatus were at rest in the ether. Since the absolute velocity of the earth does not enter into (60) it is obvious that the phase difference between the rays 1 and 2 at the point B, which is obtained by multiplying the time difference by the frequency, remains unchanged when the apparatus is rotated so as to give a different position with respect to the direction of motion of the earth. Therefore such a rotation (to a first approximation) will not cause any shift of the interference fringes.

11. The aberration of light

As we have seen in the preceding paragraph, the direction of a ray of light is, at any rate to a first approximation, independent of the absolute motion of the light source and the observer. However, the direction of a light ray depends essentially on the velocity of the light source relative to the observer. This phenomenon, which is called *aberration*, was observed in 1727 by Bradley† who noticed that the stars seem to perform a collective annual motion in the sky. This apparent motion is simply due to the fact that the observed direction of a light ray coming from a star depends on the velocity of the earth relative to the star.

In order to find the magnitude of the aberration we consider a point P' just outside the atmosphere of the earth, but in a fixed position relative to the system of reference S' following the earth. According to the above considerations, the aberration depends only on the relative velocity between the star and the observer, and we may therefore, for the sake of simplicity, assume that the star is at rest in the absolute system S. Now we consider a light ray which comes from the star and passes through the point P'. The absolute direction of the ray thus determines the direction in which the star would be observed if the earth were at rest, while the relative direction of the ray determines the apparent position of the star. Since the index of refraction n is equal to 1 at the point P', the connexion between the absolute and the relative direction of the ray is given by formula (34). Let θ and θ' be the angles between the direction of motion of the earth and the actual and the

† J. Bradley, *Phil. Trans.* **35**, 637 (1728).

apparent directions to the star, respectively. We have then $\vartheta = \pi+\theta$ and $\vartheta' = \pi+\theta'$, and from (34) we get

$$\tan\theta' = \frac{\sin\theta}{\cos\theta+v/c}. \tag{61}$$

Since now the path of the light ray from P' to the astronomer's telescope on the earth, at any rate in first approximation, is independent of the state of motion of the earth, the light will not suffer any further aberration on this path. This is true even if the ray in the course of its path passes through a strongly refracting medium as, for example, when the telescope is filled with water. Such an experiment was performed by Airy (1871),† who showed that the magnitude of the aberration was not changed by the presence of the water. The aberration formula (61) has proved to be in complete agreement with observations.

12. Michelson's experiment

As we have seen, the results of all experiments discussed up to now were in agreement with the postulate of relativity, however, the accuracy of these measurements (with the exception of Ives's experiment which was performed much later) was not good enough to allow the measurement of terms of higher than the first order. To this approximation, however, Lorentz's electron theory, which is based on the concept of an absolute ether, was in agreement with the postulate of relativity. A motion of the earth relative to the ether should, according to Lorentz's theory, influence the terms of the second order only. In order to obtain a decisive experimental test of the postulate of relativity it was, therefore, of the utmost importance to devise an experimental arrangement permitting a measurement of quantities of the second order.

This was accomplished in 1881 by Michelson,‡ who measured the velocity of light by means of the interferometer arrangement outlined in Fig. 6. By means of a glass plate P a light beam from a light source L is divided into two rays, 1 and 2, perpendicular to one another. The transmitted ray 1 is reflected by a mirror S_1 back to P, where some of the ray is reflected further into a telescope T. Analogously, the ray 2 is reflected by a mirror S_2 back to P and part of the ray goes through the glass plate and enters the telescope where it interferes with the ray 1. Even if the apparatus were at rest in the ether we should obviously observe a set of interference fringes in the telescope.

† G. B Airy, *Proc. Roy. Soc. London*, A, **20**, 35 (1871), **21**, 121 (1873); *Phil. Mag.* **43**, 310 (1872).

‡ A. A. Michelson, *Amer. Journ. of Science* (3), **22**, 20 (1881); A. A. Michelson and E. W. Morley, ibid. **34**, 333 (1887).

Now let us assume that the apparatus is placed in such a way that the path PS_1 is parallel to the direction of motion of the earth in the ether, and let the paths PS_1 and PS_2 have the same length l. By means of (35)

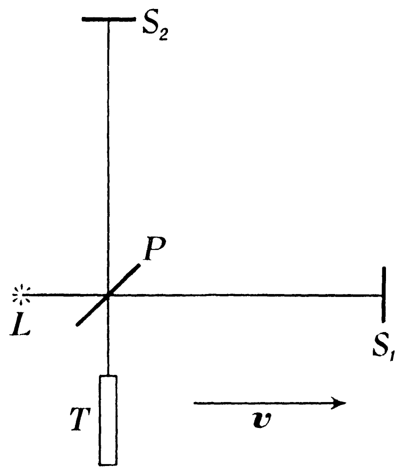

FIG. 6.

it is then easy to calculate the phase difference ΔF between the rays 1 and 2 which is due to the motion of the apparatus in the ether. For the time t_1 which the ray 1 takes to travel from P to S_1 and back again we first obtain

$$t_1 = \frac{l}{c-v}+\frac{l}{c+v} = \frac{2l}{c}\frac{1}{1-v^2/c^2}, \tag{62}$$

since $\mathbf{e}'$ on the way forth and back is parallel to the vector $\mathbf{v}$. In the same way we obtain the time t_2 which ray 2 takes to travel from P to S_2 and back and, since in this case $\mathbf{e}'$ is perpendicular to $\mathbf{v}$ along the whole path, we get by means of (35)

$$t_2 = 2l(c^2-v^2)^{-\frac{1}{2}}. \tag{63}$$

Neglecting terms of order higher than the second, we therefore get for the mentioned phase difference

$$\Delta F = \nu(t_1-t_2) = \nu l v^2/c^3. \tag{64}$$

When the apparatus is rotated through an angle of 90°, so that the path PS_2 now becomes parallel to $\mathbf{v}$, the difference in phase will be $-\Delta F$. Such a rotation of the apparatus should therefore cause a shift of the interference fringes corresponding to a change in phase of $2\Delta F$.

When the distance between the interference fringes is used as unit of length, the phase difference $2\Delta F$ gives directly the shift of the interference fringes by such a rotation of the apparatus. In Michelson's experiment, $2\Delta F$ was about $\frac{1}{3}$, so that one would expect a shift in the position of the fringes of about $\frac{1}{3}$ of the distance between the fringes. Despite the fact that Michelson would have been able to detect with certainty a shift a hundred times smaller, he could not find any effect at all. Thus for the first time we are confronted with an experiment indicating that the principle of relativity is true with an accuracy of at least second order.†

13. The contraction hypothesis

The result of Michelson's experiment meant a very serious difficulty for the ether hypothesis. Michelson himself tried to explain the absence of the effect by assuming that the ether was carried along by the earth during its motion round the sun. In this case there would be no ether wind at the surface of the earth, but perhaps at high altitudes. Michelson therefore repeated his experiment on a high mountain, but still without any detectable effect. The assumption that the ether should be completely dragged along by the earth is also in conflict with all optical experience and with Lorentz's electron theory, according to which the dragging is only partial inside the refracting media.

In order to explain the absence of any effect due to the motion of the earth in Michelson's experiment, Lorentz‡ and FitzGerald§ independently put forward the hypothesis that *any rigid body moving with velocity* $\mathbf{v}$ *is contracted in its direction of motion*, the relative contraction being equal to $(1-v^2/c^2)^{\frac{1}{2}}$. The length of the path PS_1 in Michelson's experiment (Fig. 6) would then not be l, but $l(1-v^2/c^2)^{\frac{1}{2}}$, while the length of PS_2 is unchanged, since PS_2 is at right angles to the direction of motion of the apparatus. For the time t_1 we then obtain, instead of (62),

$$t_1 = 2l(c^2-v^2)^{-\frac{1}{2}} = t_2, \tag{65}$$

t_2 being given by (63). In this case the phase difference ΔF becomes zero in agreement with Michelson's experiment.

According to this strange hypothesis, a stick which has the length l_0 when it is perpendicular to the direction of motion of the earth should obtain the shorter length

$$l = l_0(1-v^2/c^2)^{\frac{1}{2}} \tag{66}$$

† Michelson's results have been confirmed later by several investigators. See, for example, R. J. Kennedy, *Proc. Nat. Acad.* **12,** 621 (1926), and K. K. Illingworth, *Phys. Rev.* **30,** 692 (1927). Contrary to these results Miller obtained a small effect. See D. C. Miller, *Rev. Mod. Phys.* **5,** 203 (1933).

‡ H. A. Lorentz, *Amst. Verh., Akad. v. Wet.* **1,** 74 (1892).

§ G. F. FitzGerald, see O. Lodge, *London Transact.* (A) **184,** 727 (1893), in particular p. 749.

when it is turned so as to be parallel to the direction of motion of the earth. Certainly on the earth it will never be possible directly to measure this shortening, since all bodies, thus also the measuring sticks, are shortened at equal rates. An observer at rest in the ether outside the earth would, however, in principle be able to observe the shortening and he would find the earth and all objects on the earth contracted in the direction of motion of the earth.

The contraction hypothesis looks rather startling at first sight, but, as stressed by Lorentz,† it is impossible to escape from it as long as the conception of an absolute unmovable ether is maintained. For, from this view-point, the result of Michelson's experiment can directly be taken as a proof of the contraction with the same justification as the shift of the interference fringes when certain parts of the apparatus are heated is taken as a proof of a change in length of the heated parts.

In order to make the hypothesis somewhat more acceptable, Lorentz made an attempt at explaining the contraction phenomenon on the basis of the electron theory. He actually succeeded in giving a plausible explanation of the formula (66). Assuming that the material bodies are built up of electrical particles which are held together exclusively by means of electric forces, he was able to show that the equilibrium positions of the electrical particles in such a purely electrical system are changed in agreement with (66), when the system as a whole is given a constant velocity in the ether. The difficulty was only that the presupposition that the particles are held together *exclusively* by electric forces could scarcely be assumed to be satisfied in the real substances. In particular it was difficult to imagine how the charge of a single electron could be held together, unless strong attractive forces of non-electrical nature were active inside the electron. If one therefore assumes that the contraction formula (66) is valid also for a single electron, as was actually assumed by Lorentz, this must be regarded as a pure hypothesis which cannot be based on the principles of the electron theory alone. The Lorentz contraction therefore seemed to be a basic and universal phenomenon underlying the general laws of nature.

14. Validity of the principle of relativity for all physical phenomena

Michelson's experiment was only the first of a long series of attempts to determine the motion of the earth relative to the ether. These experiments include both optical and purely electromagnetic arrangements,

† See ref., p. 21.

and in each case the result was completely negative. All phenomena appeared to be independent of the motion of the earth. At the end it was impossible to doubt that *the principle of relativity is valid exactly*, not only for the mechanical phenomena, but also for all optical and electromagnetic phenomena.

We shall not here enter into a detailed discussion of all these experiments, but confine ourselves to recalling Ives's experiment mentioned in § 5, which showed that the Doppler effect also to a second approximation depends only on the relative velocity of the source of light relative to the observer, in agreement with the principle of relativity. This fact as well as the results of some of the other experiments mentioned above cannot be explained on the basis of the ether theory, even if the contraction hypothesis is added to it; for the formula (22) does not contain any quantity which has the dimension of a length. Consequently a hypothesis regarding the contraction of distances alone will not change the formula (22) into the formula (II. 90) verified by Ives's experiment.

Next, Lorentz† investigated the problem of which hypotheses must be introduced beside the contraction hypothesis so as to make all predictions of the ether theory in accordance with the principle of relativity now verified by the experiments. He found that it was necessary in every inertial system to use a special time, the so-called local time, which is different from the time in the absolute ether system. According to the contraction hypothesis, the length of a metre stick depends on the absolute velocity of the inertial system considered. Similarly, the rate of clocks and therefore also the unit of time should, according to this new hypothesis, depend on the state of motion of the inertial system. If the basic equations of the electron theory in a moving system of inertia are written in terms of these local time and space variables, they assume the same form in any system of inertia. All electromagnetic phenomena should therefore appear to be independent of the state of motion of the frame of reference. In this way it was for some time made possible to maintain the concept of an absolute ether by the introduction of new hypotheses, until Einstein (1905) realized that the very foundations of the ether theory were seriously shaken by the results of the abovementioned experiments.

† See ref., p. 21.

II

RELATIVISTIC KINEMATICS

15. Simultaneity of events

THE fruitless attempts to find any influence of the motion of the earth on mechanical, optical, and electromagnetic phenomena gave rise to the conviction among physicists that the principle of relativity was valid for all physical phenomena. This obviously changes the whole basis of our description of nature, for, as already mentioned in § 2, the concept of an absolute ether system loses its physical meaning as soon as the universal validity of the principle of relativity is accepted. All physical phenomena will then take the same course of development in any system of inertia and one can never by any physical experiment decide which system is the absolute one. All systems of inertia then become completely equivalent and it must be required of a satisfactory theory that all systems of inertia are treated on the same footing.

Einstein was the first to formulate this new standpoint and to draw the consequences of it in his fundamental paper of 1905.† We have already mentioned one of these consequences in § 2. Since the fundamental equations of electrodynamics—Maxwell's equations—must hold now in any system of inertia, it follows that the velocity of propagation of light *in vacuo* must have the same constant value $c = 3\times 10^{10}$ cm./sec. in every system of inertia. This is, of course, in conflict with the kinematical concepts derived from our usual experience and which are expressed in the addition theorem (I. 3). Thus the new experiences gained from extremely accurate experiments, which are expressed in the principle of relativity, compel us to make a revision of these kinematical concepts which by habit had obtained an *a priori* validity in the minds of physicists. Now Einstein could show that a closer analysis of the concept of velocity, i.e. a discussion of the methods by which a measurement of velocities can actually be performed, opens up the possibility of an unambiguous description of physical phenomena in accordance with the principle of relativity.

Let us consider a light signal which travels in a straight line from a point A to another point B in a given system of inertia. The velocity with which the light signal has travelled from A to B is then defined as the ratio between the distance from A to B and the time which the light

† A. Einstein, *Ann. d. Phys.* **17**, 891 (1905); *Jahrb. d. Radioaktivitat und Elektronik*, **4**, 411 (1907).

needs to travel from A to B. The measurement of the distance does not involve any difficulties, but on second thoughts it becomes clear that the measurement of the difference in time between the emission of the light signal from A and its arrival at B is not so simple. If we imagine the time of emission t_1 to be read on a clock placed at A, while the time of arrival t_2 is read on another clock at B, the difference t_2-t_1 obtained in this way will only give the real time which the light has taken to travel from A to B if the clocks at A and B are put right. This obviously requires that the hands of both clocks *simultaneously* are in the same position. But how can we make sure that two events occurring in two different places are simultaneous?

It is possible to think of various methods to synchronize the clocks at A and B. We can, for example, carry a third clock, which is set according to the clock at A, from A to B and adjust the clock at B according to it, or we can use a time signal which is sent from A to B. We shall start by a consideration of the latter method which, in practice, has proved to be the most accurate. Let us imagine that the time signal is emitted from A when the clock at A records zero. In everyday life one would then usually set the clock at B to zero when the time signal arrives at B. This is, of course, not quite correct since the time signal is propagated with a finite velocity. If we wish to be quite accurate, the clock has on the arrival of the time signal at B to be put to l/u, where l is the distance from A to B, and u is the velocity of the time signal. In order to be able to make this correction it is thus necessary to know the velocity of the time signal, but a measurement of a velocity presupposes, as shown above, that two clocks in different places are synchronized, which was just the problem which the time signal should help us to solve. Similar observations are true when we set the clocks at A and B by transporting a third clock from A to B. In this case, the transported clock has to be corrected for the influence which the transportation might possibly have had on the motion of the clock, but to find experimentally this influence it is obviously necessary in advance to dispose of two clocks in different places of which we know that they are synchronized. Here again we are moving in a circle.

All methods for the regulation of clocks meet with the same fundamental difficulty. The concept of simultaneity between two events in different places obviously has no exact objective meaning at all, since we cannot give any experimental method by which this simultaneity could be ascertained. The same is therefore true also for the concept of velocity. As stressed by Einstein, we must first *define* what we understand by simultaneity. As to definitions of concepts we are, however, to some extent free and, as we shall see in the following, it is possible to use such a definition of simultaneity that the velocity of light is constantly equal to c in all inertial systems.

16. The relativity of simultaneity

Let us imagine that we are placed in an arbitrary inertial system I and that we are provided with a large number of clocks (standard clocks) which show the same rate when placed at rest at the same place. These clocks shall be distributed at all places in I where we want to make time measurements. In order to synchronize the clocks we shall use light signals since we know experimentally a good deal about the propagation of light. Fizeau's experiment and similar more accurate measurements show, for instance, that the time which a light ray takes to traverse a closed polygon is equal to the ratio between the total length of the polygon and the universal constant c occurring in Maxwell's equations. This time can be measured by a *single* clock placed at a fixed point of the polygon, independently of the definition of simultaneity, and all distances may be measured by means of standard measuring-sticks at rest in I.

Now we choose an arbitrary point O as regulating centre, and in order to synchronize the clocks at different places in I, a light signal is emitted from O in all directions. Let the clock at O at the start of the signal record the time t_0; when this signal arrives at an arbitrary point P the clock is put to t_0+l_0/c, where l_0 is the distance from O to P measured with standard measuring-sticks at rest in I. In this way all the clocks in the inertial system I are set in a definite way. Two events occurring at two arbitrary points P and P_1 are said to be simultaneous when the clocks at P and P_1 record the same time at the moment when the events occur. Such a definition of simultaneity is completely justified if it can be shown that it does not contain any inconsistencies. In this connexion two conditions must be satisfied, viz.

1. A signal starting from O τ seconds later than the regulation signal, i.e. at the time $t_0+\tau$, shall arrive at P τ seconds later, i.e. when the clock at P records the time $t_0+l_0/c+\tau$. This condition means that the method of regulation of the clocks shall be independent of the time when the regulation is made.
2. The method shall be independent of the choice of the point which is taken as regulating centre.

The first condition is no doubt fulfilled, since all points in an inertial system are equivalent, so that two standard clocks which have the same rate when placed together at O will also have the same rate when they are installed at different points O and P.

As regards condition 2, we have only to show that a light signal emitted from an arbitrary point P_1, when the clock there records t_1, will arrive at another arbitrary point P when the clock at this point records the time

$$t = t_1 + l/c, \tag{1}$$

where l is the distance between P_1 and P. In order to prove the equation (1) we may, for simplicity, assume that the time of departure t_1 of the signal coincides with the arrival at P_1 of the regulation signal from O, i.e.

$$t_1 = t_0 + l_1/c. \tag{2}$$

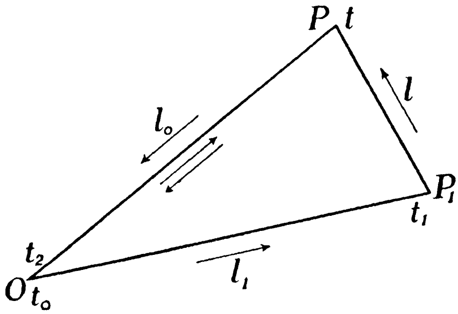

FIG. 7.

If the signal from P_1 immediately after its arrival at P is sent on to the point O, its time of arrival t_2 at O is, according to the experimental results mentioned above (Fizeau's experiment), given by

$$t_2 = t_0 + (l_1 + l + l_0)/c, \tag{3}$$

since the light signal has actually traversed a triangular path of total length $l_1 + l + l_0$ (see Fig. 7).

The clock at P showed the time t when the signal from P_1 arrived at P. When the regulation signal from O arrived at P, the clock was put to $t_0 + l_0/c$. If this signal had been reflected back to O it would, according to Fizeau's experiment, have arrived there when the clock at O recorded the time $t_0 + 2l_0/c$. Now the signal from P_1 is, however, first sent down to O $\tau = t - (t_0 + l_0/c)$ seconds later. According to the above assumption 1 it therefore arrives at O when the clock there records the time

$$t_2 = t_0 + 2l_0/c + \tau = t + l_0/c. \tag{4}$$

This time is identical with the time t_2 given in (3). From (3), (4), and (2) we now obtain

$$t = t_0+(l_1+l)/c = t_1+l/c.$$

This is just the equation (1) which we wanted to prove.

Thus the regulation method used is independent of the point chosen as regulating centre and in this way we have established a definite way of recording successive events in the inertial system I. Any event occurring at a point P at the moment when the standard clock at this place shows the time t is simply said to occur at the time t. Then also the concept of velocity assumes an exact meaning and, especially for the velocity of light, we obviously obtain the value c in all directions. For a light signal emitted from an arbitrary point P_1 at an arbitrary time t_1 has just been shown to arrive at another arbitrary point P at the time t_1+l/c, where l is the distance between P and P_1.

Let us now consider another arbitrary inertial system I'. Suppose that in this system also we place a great number of standard clocks of identical construction to those used in I. These clocks are distributed at the different points of I' and regulated in the same way as the clocks in I by means of light signals emitted from an arbitrary point O' in I'. All distances in I' are now supposed to be measured with standard measuring-sticks at rest in I'. The measuring-sticks shall be of the same type as those used in I, which means that they have the same length when brought to rest relative to each other. Since, according to the principle of relativity, Fizeau's experiment gives the same result in I' as in I, it is clear that this method of synchronizing the clocks in I' provides a consistent time description.

When the distances and the time differences are measured with clocks and measuring-sticks in I', it then follows in the same way as in I that the velocity of light relative to I' is also constant and equal to c in all directions. An event occurring at a given point P' of I' at the moment when the clock at this point registers the time t' is said to occur at the time t' relative to I'. In general this time will be different from the time t at which the same event occurs relative to I. Two events occurring at different points are now, of course, called simultaneous relative to I' if they occur at the same time t' in I'.

Thus the concept of simultaneity has lost its absolute meaning since two events occurring simultaneously for observers in I generally will not be simultaneous for observers in I'. Let us, for instance, consider two events occurring at two points A and B which are fixed in I. Since the velocity of light is c in all directions, the criterion for these events to be simultaneous relative to I is obviously that two light signals emitted from A and B at the moment when the events occur shall meet in the centre C of the line connecting A and B. A similar criterion for simultaneity is true also relative to I'. Now let the two events be simultaneous relative to I and let us, for instance, imagine that the line connecting A and B is parallel to the direction of the velocity $\mathbf{v}$ of I' relative to I. Then consider the two points A' and B' in I' which, at the moment when the events occur, coincide with the points A and B. Simultaneously (relative to I) the centre C' between A' and B' will coincide with C. Since now C', just as A' and B', moves together with I' with a velocity $\mathbf{v}$ relative to I, C' will not coincide with C at the moment when the light signals from A and B meet in C. The light signals will thus *not* meet in C' and, according to the above-mentioned criterion, the two events are not simultaneous relative to I'.

The concept of simultaneity between two events in different space points consequently has an exact meaning only in relation to a given inertial system. Only in the approximation where the velocity of light can be regarded as infinitely great compared with all other velocities which occur is it permissible to speak of an absolute simultaneity independent of the states of motion of the observers. Such an approximation is quite sufficient in daily life and in many cases also in physics, and this explains the deep-rooted subjective belief in the existence of an absolute time and in absolute simultaneity.

17. The special Lorentz transformation

In a given inertial system I an event which occurs at a point P at the time t can be characterized by four figures, viz. the three coordinates specifying the point P and the time parameter t. These four figures are called the space-time coordinates of the event. If, for instance, we use a Cartesian system of coordinates in the inertial system I, the space-time coordinates of the event are $\{x, y, z, t\}$, where $\mathbf{x} = (x, y, z)$ are the Cartesian coordinates of the point P. The coordinates (x, y, z) are found by measuring the lengths of the projections of the coordinate vector $\mathbf{x}$ on the Cartesian axes by means of standard measuring-sticks at rest in the inertial system I, while the time t is read on the standard clock which is placed at rest at the point P.

In this way a definite system of space-time coordinates S is attached to the inertial system I. When S is given, the frame of reference, i.e. the inertial system I, is completely determined. On the other hand, different systems of space-time coordinates S can, of course, be used in the same frame of reference I, for example, we may use polar coordinates instead of Cartesian coordinates for the specification of the different points in space. In the special theory of relativity we shall, however, always use space-time coordinates of the above-mentioned kind, and thus need not distinguish between the frame of reference I and the coordinate system S. On the other hand, in the general theory of relativity it will appear necessary to differentiate between the frame of reference and the system of coordinates which is used for the fixation of the events occurring in the frame of reference.

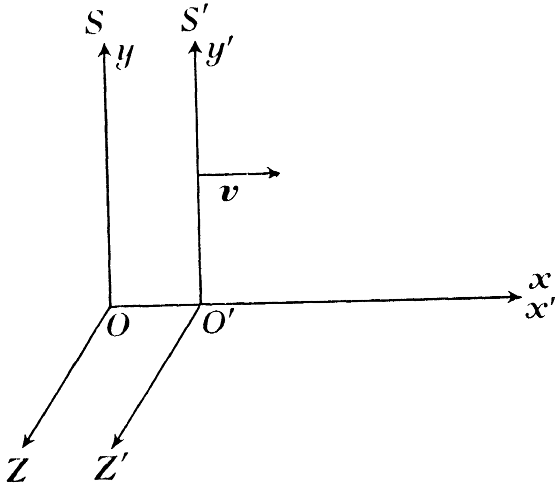

FIG. 8.

If we consider another inertial system, an event will also in this system be specified by four space-time coordinates (x', y', z', t') defining a space-time system of coordinates S'. The coordinates (x', y', z', t') are found in the same way as the coordinates in S by means of standard measuring-sticks and standard clocks now at rest in I'. Our primary task will be to find the connexion between the space-time coordinates of the same event in S and S', i.e. the transformation corresponding to the Galilean transformation (I. 1) in non-relativistic kinematics. Since any uniform translatory motion relative to S is uniform also relative to S', the variables (x', y', z', t') must obviously be linear functions of (x, y, z, t).

For convenience we shall assume that the Cartesian axes in S and S' are parallel to each other and that S' is moving relative to S with velocity $\mathbf{v}$ in the direction of the positive x-axis. Let us, moreover, assume that the origin O' of S' coincides with the origin O of S at the time

$$t = t' = 0.$$

Now consider all the points in S' which form a plane

$$y' = a' = \text{constant} \tag{5}$$

parallel to the $x'z'$-plane. These points will also form a plane

$$y = a = \text{constant} \tag{6}$$

in S parallel to the xz-plane. The constants a' and a denote the distances between these planes and the xz-plane, and, since these distances are measured by means of measuring-sticks in different states of motion, the ratio

$$\kappa = a'/a \tag{7}$$

might turn out to be different from 1.

The ratio κ can, however, only depend on the relative velocity v and a simple relativity argument shows that κ actually must be equal to 1. For, if we change the signs on the x- and z-axes as well as on the x'- and z'-axes, neither a nor a' is changed, but now the two inertial systems have exchanged roles, S now moving with the velocity v relative to S' in the direction of the positive x'-axis. Hence, we may conclude, just as before, that

$$\kappa = a/a'. \tag{8}$$

From (7) and (8) it follows, however, that $\kappa^2 = 1$ and, since the positive directions of the y- and the y'-axes are the same, a and a' must have the same sign. Hence,

$$\kappa = 1, \qquad a' = a. \tag{9}$$

These considerations show that an event occurring at a point with the coordinate y in S will have a coordinate y' with respect to S', which is given by

$$y' = y. \tag{10}$$

In the same way we find that the z-coordinate is transformed according to the equation

$$z' = z. \tag{11}$$

In order to find the transformation equations for the two other space-time coordinates we make use of the fact that a light signal in S and in S' is propagated in all directions with the same velocity c. If the light signal starts from the coinciding points O and O' at the time $t = t' = 0$, the propagation of the spherical light wave is described in S by the equation

$$x^2+y^2+z^2-c^2t^2 = 0. \tag{12}$$

In S' the wave is analogously given by the equation

$$x'^2+y'^2+z'^2-c^2t'^2 = 0. \tag{12'}$$

Now put

$$s^2 = x^2+y^2+z^2-c^2t^2 \tag{13}$$

and

$$s'^2 = x'^2+y'^2+z'^2-c^2t'^2. \tag{13'}$$

For any set of values of the variables (x, y, z, t) which makes s^2 equal to zero, s'^2 must then also be zero, and since the connexion between

$$(x', y', z', t') \quad \text{and} \quad (x, y, z, t)$$

is linear, this is only possible when s'^2 is proportional to s^2, i.e.

$$s'^2 = \kappa(v)s^2, \tag{14}$$

where κ is a constant which can depend only on the relative velocity v. By the same relativity argument which was used in connexion with equation (8) we see at once that the constant κ must actually be equal to 1. Thus (14) reduces to

$$s'^2 = s^2, \tag{15}$$

i.e. the quantity s^2 defined by (13) is an invariant.

By means of (10) and (11), (15) can be written in the form

$$x^2 - c^2t^2 = x'^2 - c^2t'^2. \tag{16}$$

Since (16) should be fulfilled for an arbitrary event, x' and t' must be linear functions of x and t alone. Therefore we can write

$$\left.\begin{aligned} x' &= \alpha x + \beta t, \\ t' &= \gamma x + \delta t, \end{aligned}\right\} \tag{17}$$

where the constants α, β, γ, δ have to be determined so that (16) is satisfied for all x and t.

For the origin O' we have $x' = 0$. Thus we obtain from the first equation (17) for the motion of O' relative to S

$$x = -\beta t/\alpha,$$

and, since the velocity of O' relative to S is v, it follows that

$$\beta = -\alpha v. \tag{18}$$

For the origin O we have, however, $x = 0$. By introducing $x = 0$ into the equations (17) we get, after elimination of t, the equation

$$x' = \beta t'/\delta$$

describing the motion of O relative to S'. For symmetry reasons the velocity of O relative to S' must, however, be $-v$, which together with (18) gives

$$\beta = -\delta v = -\alpha v, \quad \text{i.e. } \delta = \alpha. \tag{19}$$

By means of (18) and (19) equation (17) can be written in the form

$$\left.\begin{aligned} x' &= \alpha(x - vt), \\ t' &= \gamma x + \alpha t. \end{aligned}\right\} \tag{20}$$

Introduction of these equations into (16) gives

$$x^2 - c^2t^2 = (\alpha^2 - \gamma^2c^2)x^2 - \alpha^2(1 - v^2/c^2)c^2t^2 - 2\alpha(\gamma c^2 + \alpha v)xt. \tag{21}$$

Since this latter equation will be satified by all possible values of the independent variables x and t, the coefficient of the variables x^2, t^2, and xt, respectively, on both sides of the equation must be equal. This gives three equations for the determination of the two quantities α and γ. The last two of these equations give immediately

$$\alpha = (1-v^2/c^2)^{-\frac{1}{2}} \tag{22}$$

and

$$\gamma = -\alpha v/c^2 = -\frac{v/c^2}{\sqrt{(1-v^2/c^2)}}. \tag{23}$$

The remaining equation, expressing the fact that the coefficients of x^2 on both sides of the equation are equal, is then identically fulfilled, which shows that equation (19), i.e. the assumption that O relative to S' moves with a velocity $-v$, is in accordance with equation (16).

From (10), (11), (20), (22), and (23) we finally get the following transformation equations for the space-time coordinates of an arbitrary event

$$\left.\begin{aligned} x' &= \frac{x-vt}{\sqrt{(1-v^2/c^2)}}, \quad y' = y, \quad z' = z \\ t' &= \frac{t-vx/c^2}{\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}. \tag{24}$$

The inverse relations which are obtained by solving the four equations (24) with respect to the variables x, y, z, t are

$$\left.\begin{aligned} x &= \frac{x'+vt'}{\sqrt{(1-v^2/c^2)}}, \quad y = y', \quad z = z' \\ t &= \frac{t'+vx'/c^2}{\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}. \tag{24'}$$

They may be obtained from (24) by interchanging the primed and the unprimed variables and replacing v by $-v$.

In (18) the quantity v was introduced as the velocity of O' relative to S. It follows from (24), however, that any fixed point P' in S' with constant values of coordinates x', y', z' moves with velocity v relative to S in the direction of the x-axis. Analogously, we see from (24′) that each fixed point P in S moves with velocity $-v$ relative to S' in the direction of the x'-axis. The quantity v therefore denotes simply the relative velocity of the two systems of inertia.

Lorentz was the first to introduce the transformation equations (24) and (24′) and, therefore, they are usually called Lorentz transformations. The derivation of these equations from the point of view of the principle of relativity is, however, due to Einstein.† In view of the special position of the Cartesian axes in Fig. 8, we here speak of a *special* Lorentz transformation. Under such a transformation the quantity s^2 defined by (13) is invariant. If we put $c = \infty$, (24) goes over into the special Galilean transformation (I. 2).

18. The most general Lorentz transformation

The transformation of the space-time coordinates in the more general case, where the relative velocity of S' and S is not parallel to the x-axis and where the rectangular coordinates in S and S' have arbitrary orientations relative to each other, can obviously be obtained by means of a suitable combination of spatial rotations of the axes in S' and S together with a special Lorentz transformation (24). Since s^2 is unchanged by spatial rotations, s^2 is thus also invariant by these more general Lorentz transformations.

Sometimes we shall need explicit expressions for the Lorentz transformations in the general case and it is convenient then to use the following vector representation. Let us, for a moment, again consider a special Lorentz transformation corresponding to the orientation of the coordinates shown in Fig. 8. We can now depict the space vectors

$$\mathbf{x} = (x, y, z) \quad \text{and} \quad \mathbf{x}' = (x', y', z')$$

of the space points of the event considered in S and S', respectively, *in one and the same* vector space, (x, y, z) and (x', y', z') being regarded as the components of the image vectors in a *fixed* system of coordinates in this abstract three-dimensional vector space. In the same vector space the vector $\mathbf{v}$, representing the velocity of the system S' relative to S, is depicted as a vector with components $(v, 0, 0)$. The special Lorentz transformation (24) can then obviously be written as a relation between the image vectors $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{v}$ and the variables t and t' of the form

$$\left.\begin{aligned} \mathbf{x}' &= \mathbf{x}+\mathbf{v}\left[\frac{(\mathbf{x}.\mathbf{v})}{v^2}\{(1-v^2/c^2)^{-\frac{1}{2}}-1\}-t(1-v^2/c^2)^{-\frac{1}{2}}\right] \\ t' &= (1-v^2/c^2)^{-\frac{1}{2}}\{t-(\mathbf{v}.\mathbf{x})/c^2\} \end{aligned}\right\}, \tag{25}$$

where $(\mathbf{v}.\mathbf{x}) = v_x x+v_y y+v_z z$ is the scalar product of the image vectors $\mathbf{x}$ and $\mathbf{v}$. It is immediately seen that the three components of the vector equation in (25) are identical with the first three equations (24).

Introducing in the same manner the image vector $\mathbf{v}' = (-v, 0, 0)$ which represents the velocity of the system S relative to S', the inverse

† Cf., for example, A. Einstein, *Über die spezielle und die allgemeine Relativitatstheorie*, Braunschweig 1917.

equations (24′) can obviously be written in the form

$$\left.\begin{aligned} \mathbf{x} &= \mathbf{x}'+\mathbf{v}'\left[\frac{(\mathbf{x}'.\mathbf{v}')}{v^2}\{(1-v^2/c^2)^{-\frac{1}{2}}-1\}-t'(1-v^2/c^2)^{-\frac{1}{2}}\right] \\ t &= (1-v^2/c^2)^{-\frac{1}{2}}\{t'-(\mathbf{v}'.\mathbf{x}')/c^2\} \end{aligned}\right\}. \quad (25')$$

Since

$$\mathbf{v}' = -\mathbf{v} \quad (26)$$

the inverse equations (25′) are in this case obtained from (25) by interchanging $(\mathbf{x}',t')$ and $(\mathbf{x},t)$ and replacing $\mathbf{v}$ by $-\mathbf{v}$.

By a rotation of the Cartesian axes in S the components of $\mathbf{x}$ and $\mathbf{v}$ undergo an orthogonal transformation, and since the axes in the abstract vector space are supposed to be fixed, this means that the rotation of the Cartesian axes in S induces a corresponding inverse rotation of the image vectors $\mathbf{x}$ and $\mathbf{v}$, while the vectors $\mathbf{x}'$ and $\mathbf{v}'$ remain unchanged. In a similar way, a rotation of the Cartesian axes in S' induces the corresponding inverse rotation of the image vectors $\mathbf{x}'$ and $\mathbf{v}'$.

Now let us first consider the case where the Cartesian axes in S and S' are subjected to the *same* rotations (starting from their position as shown in Fig. 8). This means that the sets of variables (x,y,z) and (x',y',z') are subjected to orthogonal transformations with the same coefficients. Then the image vectors $\mathbf{x}$, $\mathbf{v}$, $\mathbf{x}'$, $\mathbf{v}'$ also suffer the same rotations and therefore the relations between these image vectors will still be given by (25), (25′), and (26). In this case we speak of a Lorentz transformation without rotation, since the angles (measured in S and S', respectively) through which the Cartesian axes in S and S' should be turned in order to obtain the orientation shown in Fig. 8 are the same, so that in a certain sense the Cartesian axes in S and S' have the same orientation (cf., however, the considerations in § 19).

If v_x, v_y, v_z denote the components of the velocity of the system S' relative to S, and if we write $\gamma = (1-v^2/c^2)^{-\frac{1}{2}}$, the vector equations (25) are only a short way of writing the four equations

$$\left.\begin{aligned} x' &= \{1+(\gamma-1)v_x^2/v^2\}x+(\gamma-1)v_x v_y y/v^2+(\gamma-1)v_x v_z z/v^2-v_x\gamma t \\ y' &= (\gamma-1)v_y v_x x/v^2+\{1+(\gamma-1)v_y^2/v^2\}y+(\gamma-1)v_y v_z z/v^2-v_y\gamma t \\ z' &= (\gamma-1)v_z v_x x/v^2+(\gamma-1)v_z v_y y/v^2+\{1+(\gamma-1)v_z^2/v^2\}z-v_z\gamma t \\ t' &= -\gamma v_x x/c^2-\gamma v_y y/c^2-\gamma v_z z/c^2+\gamma t \end{aligned}\right\}, \quad (27)$$

which thus represent a general Lorentz transformation without rotation. On account of (26) the inverse equations (25′) are again obtained from (27) by interchanging the variables (x,y,z,t) and (x',y',z',t') and substituting $(-v_x, -v_y, -v_z)$ for (v_x, v_y, v_z), respectively.

Proceeding now to the consideration of the case where the Cartesian axes in S and S' do not have the same orientation, we must keep in mind that the Cartesian axes in S and S' must be subjected to *different* rotations in order to attain the orientation of the axes shown in Fig. 8. While the last equation (25) remains valid without change, the first equation has to be replaced by

$$\mathfrak{D}\mathbf{x}' = \mathbf{x}+\mathbf{v}\{(\gamma-1)(\mathbf{x}.\mathbf{v})/v^2-\gamma t\}, \tag{28a}$$

where $\mathfrak{D}$ is the rotation operator which transforms the image vector $\mathbf{x}'$ into the vector $\mathfrak{D}\mathbf{x}'$ corresponding to a Lorentz transformation without rotation. Thus $\mathfrak{D}^{-1}$ represents the rotation of the Cartesian axes in S', which would give these axes the same orientation (in the above-mentioned sense) as the axes in S. Instead of the equation (26) we have

$$\mathfrak{D}\mathbf{v}' = -\mathbf{v}, \tag{29}$$

i.e. the components (v'_x, v'_y, v'_z) of the velocity of S relative to S' are in this case not equal to $(-v_x, -v_y, -v_z)$. Multiplying (28 a) by the inverse rotation operator $\mathfrak{D}^{-1}$ and applying (29), the Lorentz transformation may in this general case be written in the form

$$\left.\begin{aligned} \mathbf{x}' &= \mathfrak{D}^{-1}\mathbf{x}-\mathbf{v}'\{(\gamma-1)(\mathbf{x}.\mathbf{v})/v^2-\gamma t\} \\ t' &= \gamma\{t-(\mathbf{v}.\mathbf{x})/c^2\} \end{aligned}\right\}. \tag{28b}$$

$\mathfrak{D}$ can also be interpreted as the rotation which has to be applied to the axes in S in order to obtain the same orientation of the axes in S and S'. The relations inverse to (28 b) are therefore

$$\left.\begin{aligned} \mathbf{x} &= \mathfrak{D}\mathbf{x}'-\mathbf{v}\{(\gamma-1)(\mathbf{x}'.\mathbf{v}')/v^2-\gamma t'\} \\ t &= \gamma\{t'-(\mathbf{v}'.\mathbf{x}')/c^2\}, \quad v' = |\mathbf{v}'| = |\mathbf{v}| = v \end{aligned}\right\}, \tag{28'}$$

which also can be proved directly by introducing (28′) into the right-hand side of (28 b).

It is easily seen that the equations (25) and (28) satisfy the equation (15) which can be written in the vector form

$$(\mathbf{x}.\mathbf{x})-c^2t^2 = (\mathbf{x}'.\mathbf{x}')-c^2t'^2. \tag{30}$$

If we put $c = \infty$, (25) becomes the general Galilean transformation (I. 1).

Until now we have assumed that the origins O and O' coincide at the time $t = t' = 0$ and, accordingly, the Lorentz transformations are *homogeneous* transformations of the space-time coordinates. We shall now abandon this assumption and consider a displacement of the origin of the space and time coordinates in S'. This means that we have to replace (x', y', z', t') in (24), (25), and (28) by $x'-x'_0$, $y'-y'_0$, $z'-z'_0$, $t'-t'_0$, respectively, where x'_0, y'_0, z'_0, t'_0 are constants. By an inhomogeneous Lorentz transformation of this type the quantity s^2 will not be invariant any more. However, if we consider two events with the coordinates x_1, y_1, z_1, t_1 and x_2, y_2, z_2, t_2, respectively, the differences

$$\Delta x = x_1 - x_2, \qquad \Delta y = y_1 - y_2, \qquad \ldots$$

between the coordinates of the two events will also be transformed for an *inhomogeneous* Lorentz transformation according to the equations (25), (28) since the constants x'_0, y'_0, z'_0, t'_0 will disappear when the differences are formed. Therefore the quantity Δs^2, defined by the equation

$$\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2 \Delta t^2, \tag{31}$$

will be invariant by an arbitrary inhomogeneous Lorentz transformation.

According to the principle of relativity all physical phenomena have the same course of development in all systems of inertia and it must be required that, in the theoretical description of the phenomena, all inertial systems are treated on the same footing, i.e. the fundamental equations of physics must have the same form in every inertial system. In other words the fundamental equations must be *form-invariant* or *covariant* under Lorentz transformations. This requirement, which is the formal expression of the principle of relativity, has proved to be very useful in the development of new theories.

As we shall see below, this requirement of form-invariance is automatically fulfilled for Maxwell's fundamental equations of electrodynamics *in vacuo*. On the other hand, Newton's fundamental equations of mechanics do not satisfy this requirement, since these equations, as shown on pp. 3–4, are covariant under Galilean transformations. Newtonian mechanics, therefore, is valid only in the approximation where Lorentz transformations and Galilean transformations can be regarded as identical, i.e. where all the velocities occurring are small compared with the velocity of light. But for all mechanical phenomena, where velocities of the same order of magnitude as the velocity of light are involved, the Newtonian equations must be replaced by Einstein's *relativistic* equations of mechanics which are covariant under Lorentz transformations (cf. Chapter III).

19. Contraction of bodies in motion

From the Lorentz transformation (24) we can now draw certain conclusions regarding the intercomparison of measuring-sticks and clocks in the systems S and S'. Consider a measuring-rod which is at rest relative to S' and is placed parallel to the x'-axis (Fig. 8). The end-points of the rod therefore have constant coordinates x'_1 and x'_2 and the length of the rod in S' (its rest length) is

$$l^0 = x'_2 - x'_1.$$

According to the first equation (24) the motion of the two end-points relative to S is given by the equations

$$\left.\begin{aligned} x_1(t) &= vt + (1-v^2/c^2)^{\frac{1}{2}} x'_1 \\ x_2(t) &= vt + (1-v^2/c^2)^{\frac{1}{2}} x'_2 \end{aligned}\right\}. \tag{32}$$

Now it is natural to define the length l of the rod relative to S as the difference between simultaneous coordinate values of the end-points; by simultaneity in this connexion we understand simultaneity relative to S. From (32) we then obtain

$$l = x_2(t) - x_1(t) = (x'_2 - x'_1)(1-v^2/c^2)^{\frac{1}{2}} = l^0(1-v^2/c^2)^{\frac{1}{2}}, \tag{33}$$

which is independent of t. On the other hand, since the systems S and S' are completely equivalent, a metre stick at rest on the x-axis of S which has the length l^0 in S will have a length l relative to S', which again is given by (33).

A metre stick which is placed perpendicular to the x-axis will, however, according to (24), have the same length in S as in S'. We may therefore quite generally say that a body which moves with a velocity v relative to an arbitrary inertial system S is contracted in the direction of its motion according to the equation (33), while the transverse dimensions are independent of the motion of the body. If V^0 is the rest volume of the body, i.e. the volume measured in an inertial system following the body in its motion, its volume V in S is thus given by

$$V = V^0(1-v^2/c^2)^{\frac{1}{2}}. \tag{34}$$

The equation (33) is identical with the Lorentz equation (I. 66). However, as regards the physical interpretation, there is a difference in principle between the two equations. In (I. 66) l^0 was the length of the metre stick at rest in the ether while l was the length of the stick in motion with a velocity v relative to the ether. Thus, according to the original point of view of Lorentz, the metre stick is attributed an absolute length which is independent of the state of motion of the *observer*. In equation (33), however, l^0 is the rest length of the metre stick, i.e. the length measured in an inertial system following the stick, and l is the length measured in an arbitrary inertial system relative to which the stick has the velocity v. Therefore, according to relativistic conceptions, the notion of the length of a stick has an unambiguous meaning only in relation to a given inertial frame, this length being different for the different systems of inertia. This means, however, that the concept of length has lost its absolute meaning. We can only speak of an absolute length in the approximation where the velocity of light can be regarded as infinitely large.

In spite of the fact that the concept of simultaneity enters into the above definition of the length of a moving rod, equation (33), as pointed out by Einstein, can in principle be verified by experiment without the use of clocks. Let us consider two rods M_1 and M_2 with the same rest length l^0 moving along the x-axis in S with the velocities v and $-v$, respectively. Since the length of a stick according to (33) depends on the square of the velocity only, M_1 and M_2 must have the same constant length l relative to S, which means that they must coincide at a certain time t. In other words, the coincidences of the two ends of M_1 with the ends of M_2 must be simultaneous events relative to S. Let these coincidences occur at the points A and B in S. Then a subsequent measurement of the distance AB with a standard metre stick in S gives the magnitude of l.

Even if it is, of course, impossible to perform such an experiment in practice with an accuracy sufficiently high to verify equation (33), this consideration shows that the Lorentz contraction given by (33) is a real effect observable in principle by experiment. It expresses, however, not so much a quality of the moving stick itself as rather a reciprocal relation between measuring-sticks in motion relative to each other. In this connexion it is natural to ask for the cause of the contraction. According to the principle of relativity, the answer must be that such a question is just as delusive as if, after the discovery of the law of inertia, the question were put why a body left to itself will continue to move straight forward with uniform velocity. While such a question was well justified in Aristotelian physics it must be rejected as meaningless after Galileo's discovery. According to Galilean and Newtonian mechanics only the deviations from uniform translatory motions require a cause.

While Lorentz attempted to explain the contraction phenomenon on the basis of the electron theory, Einstein's deduction of equation (33), based on the principle of relativity alone, shows that the contraction phenomenon is of a much more fundamental character. Instead of considering the contraction to be a phenomenon which has to be *explained* on the basis of an atomistic theory of material bodies, it should rather be regarded as something *elementary* which cannot be traced back to simpler phenomena. Actually it represents a requirement which must be satisfied by any atomistic theory, viz. the requirement of covariance under Lorentz transformations. If, on the other hand, this requirement is fulfilled, the contraction of a macroscopic body in motion can obviously be deduced from the theory of the atomic structure of the body.

Before leaving this section we shall consider the somewhat more general case where the connexion between space-time coordinates is given by a Lorentz transformation without rotation. Let $\mathbf{x}'_1$ and $\mathbf{x}'_2$ be the coordinate vectors of two fixed points P'_1 and P'_2 in the system of coordinates S'. The straight line connecting these points represents a fixed vector $\mathbf{r}' = \mathbf{x}'_2 - \mathbf{x}'_1$ in S'. At the time t the points P'_1 and P'_2 will have coordinate vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in S which are obtained from the first equation (25) by putting $\mathbf{x} = \mathbf{x}_1$, $\mathbf{x}' = \mathbf{x}'_1$ and $\mathbf{x} = \mathbf{x}_2$, $\mathbf{x}' = \mathbf{x}'_2$, respectively. Then by subtraction of these equations we obtain

$$\mathbf{r}' = \mathbf{r} + \mathbf{v}\{(1-v^2/c^2)^{-\frac{1}{2}} - 1\}\frac{(\mathbf{r}\cdot\mathbf{v})}{v^2}, \tag{35}$$

where
$$\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$$

is the vector connecting two simultaneous positions of P'_1 and P'_2 in S. If $\mathbf{r} = \mathbf{r}_{\parallel} + \mathbf{r}_{\perp}$ is decomposed into its components $\mathbf{r}_{\parallel}$ and $\mathbf{r}_{\perp}$ parallel and perpendicular to $\mathbf{v}$, respectively, and analogously $\mathbf{r}' = \mathbf{r}'_{\parallel} + \mathbf{r}'_{\perp}$, (35) can also be written

$$\mathbf{r}'_{\parallel} = \mathbf{r}_{\parallel}(1-v^2/c^2)^{-\frac{1}{2}}, \qquad \mathbf{r}'_{\perp} = \mathbf{r}_{\perp},$$

which shows that it is only the parallel component which suffers a Lorentz contraction. The inverse relation to (35) is

$$\mathbf{r} = \mathbf{r}' + \mathbf{v}\frac{\mathbf{r}'\cdot\mathbf{v}}{v^2}\{(1-v^2/c^2)^{\frac{1}{2}} - 1\}, \tag{35'}$$

which can be verified by inserting (35′) in the right-hand side of (35).

From (35′) it follows that a fixed vector in S' parallel to the x'-axis is in general not parallel to the x-axis as judged by an observer in S. Even in the case of a Lorentz transformation without rotation the Cartesian axes in S' will thus from the point of view of an observer in S generally not be parallel to the axes in S. Hence if we state that the Cartesian axes have the same orientation in a Lorentz transformation without rotation this should be understood in the way described on p. 42. For two fixed vectors $\mathbf{r}'_1$ and $\mathbf{r}'_2$ in S' which satisfy the condition $(\mathbf{r}'_1\cdot\mathbf{r}'_2) = 0$ we can very well have $(\mathbf{r}_1\cdot\mathbf{r}_2) \neq 0$. Therefore the rectangular axes of coordinates in S' will generally not even be perpendicular to each other when looked upon from the system S. For this reason it was necessary to depict the vectors $\mathbf{x}$ and $\mathbf{x}'$ in the neutral abstract vector space introduced on p. 41.

20. The retardation of moving clocks. The clock paradox

Now consider a standard clock C' which is placed at rest in S' at a point on the x'-axis with the coordinate $x' = x'_1$ (Fig. 8). When the clock C' records the time $t' = t'_1$, the standard clock in S which C' is passing by at that moment will record a time t_1 given by the Lorentz transformation (24'),

$$t_1 = \gamma(t'_1+vx'_1/c^2) \quad (\gamma = (1-v^2/c^2)^{-\frac{1}{2}}).$$

Somewhat later, when C' records the time $t' = t'_2$, it coincides with another clock in S which records the time

$$t_2 = \gamma(t'_2+vx'_1/c^2).$$

By subtraction of these equations we obtain

$$\Delta t = t_2-t_1 = \gamma(t'_2-t'_1) = (1-v^2/c^2)^{-\frac{1}{2}}\Delta\tau, \tag{36}$$

i.e. a clock which is moving with the velocity v relative to S will be slow compared with the clocks in S.

When we keep in mind that the systems S and S' are equivalent it is obvious that a clock at rest in S similarly will lag behind the clocks in S'.

Since the Lorentz transformation from which this result is deduced is based upon the method of synchronizing the clocks in S and S', discussed in § 16, one could think that the retardation of moving clocks described by (36) was only apparent. However, just as in the case of the Lorentz contraction, it can be shown that equation (36) contains a statement regarding the rate of moving clocks which in principle may be verified by experiment. Consider two standard clocks C_1 and C_2 placed together at the origin of an arbitrary inertial frame of reference. At the time $t = 0$ the clock C_2 is set in uniform motion along the x-axis with the velocity v. At the time $t = t_P$ it has reached a point P on the x-axis and, according to (36), it will record the time $t_P(1-v^2/c^2)^{\frac{1}{2}}$. Immediately after arriving at P, C_2 is sent back to O with the velocity $-v$. It arrives at O when the clock C_1 records the time $t = t_1 = 2t_P$. Since, according to (36), the rate of C_2 is independent of the sign of v, C_2 will on its arrival at O record the time $t_2 = 2t_P(1-v^2/c^2)^{\frac{1}{2}}$. The clocks C_1 and C_2, which at the start of C_2 both showed the time zero, will thus, after the return of C_2, record times t_1 and t_2, respectively, connected by the equation

$$t_2 = (1-v^2/c^2)^{\frac{1}{2}}t_1. \tag{37}$$

The difference between t_2 and t_1 may now in principle be measured directly by a comparison of the readings of C_1 and C_2 before and after the whole process.

As mentioned by Einstein in his first paper on the theory of relativity, this consequence of the theory gives rise to a paradox which in the past has played an important role in the discussions on the consistency of the theory. Suppose that we introduce a frame of reference R which follows the clock C_2 during its motion from O to P and back again. Since, now, the motion of C_1 relative to R is quite analogous to the motion of C_2 relative to S, one would think that an observer in R would find that the clock C_1 is slower than C_2, in contradiction with (37). This argument is wrong, however, since equation (36) is valid only in an inertial system and therefore it is not applicable to the system R which during the change of velocity of C_2 from v to $-v$ is accelerated relative to the fixed stars. The question which equation should be used in R instead of (36) cannot be answered in the special theory of relativity which only allows treatment of the physical phenomena in frames of reference in uniform motion. This discussion clearly shows the desirability of an extension of the special theory of relativity to a general theory which allows the use of systems of coordinates in arbitrary motion. (For the final solution of the clock paradox, see Chapter VIII, § 98, p. 258.)

Now let us again consider a standard clock moving with a uniform velocity $\mathbf{u}$ relative to an inertial system S. The time recorded by the moving clock itself is called the proper time τ of the clock. According to (36) we have the following relation between the increase in proper time $d\tau$ and the increase dt of the time in S.

$$d\tau = (1-u^2/c^2)^{\frac{1}{2}}\, dt. \tag{38}$$

This equation is now assumed to be valid also for an arbitrarily moving clock where u is the momentary velocity of the clock. Hence we assume that *the acceleration of the clock relative to an inertial system has no influence on the rate of the clock, and that the increase in the proper time of the clock at any time is the same as that of the standard clocks in the rest system S^0, i.e. the system in which the clock is momentarily at rest.*

It should now also be possible to deduce the retardation of moving clocks from the fundamental laws of mechanics governing the running of the clockwork. But, just as in the case of the Lorentz contraction, it is more adequate to regard the retardation phenomenon as an elementary phenomenon which is a direct consequence of the principle of relativity. If we took Newtonian mechanics as a basis for the calculation of the working of the clock, no retardation of moving clocks would be found, since the time in Newtonian fundamental equations is an invariant parameter (cf. I. 1 b), but this just shows that the Newtonian equations are not accurate in the region where $(1-v^2/c^2)^{\frac{1}{2}}$ differs appreciably from 1. If, on the other hand, we use the exact relativistic equations of mechanics in the description of the working of the clocks (cf. Chapters III and VI), the retardation effect must of course follow as a consequence of these equations.

In view of the fact that an arbitrary physical system can be used as a clock, we see that any physical system which is moving relative to a system of inertia must have a slower course of development than the same system at rest. Consider, for instance, a radioactive process. The mean life τ of the radioactive substance, when moving with a velocity v, will thus be larger than the mean life τ^0 when the substance is at rest. From (36) we obtain immediately

$$\tau = (1-v^2/c^2)^{-\frac{1}{2}}\tau^0. \tag{39}$$

In general, v is so small relative to c that we need not discriminate between τ and τ^0. In recent years, however, very rapidly moving radioactive systems—the mesons—have been observed in cosmic radiation where the factor $(1-v^2/c^2)^{-\frac{1}{2}}$ can be of the order of magnitude of 100 or more. Therefore equation (39) has been of essential importance in the interpretation of the phenomena connected with the decay of the mesons.

We may also use a radiating atom as clock, the number of light waves emitted per second being a measure of the rate of the atomic clock. If ν^0 is the proper frequency of the atom, i.e. the frequency of the emitted light measured in the system of inertia S^0 in which the atom is at rest, the number of waves emitted per unit time in this system is just ν^0. In the system S relative to which the atom is moving with the velocity v, the number of waves emitted per unit time will then be $\nu^0(1-v^2/c^2)^{\frac{1}{2}}$, for, according to (36), a unit time interval in S corresponds to the time interval $\Delta\tau = (1-v^2/c^2)^{\frac{1}{2}}$ in the rest system S^0. If now the moving atom has no radial velocity relative to an observer in S, this number will be equal to the frequency ν found by the observer in S, the number of emitted waves being equal to the number of waves arriving per time unit. Consequently we have

$$\nu = (1-v^2/c^2)^{\frac{1}{2}}\nu^0 \tag{40}$$

when the radial velocity is zero. This is, for example, the case when the atom moves with the velocity v along a circle whose centre is the observer or when the direction of the light arriving at the observer is perpendicular to the direction of motion of the atom.

According to the theory of relativity we must therefore expect a shift in frequency also for perpendicularly incident light, viz. a shift towards smaller frequencies, in contradistinction to the non-relativistic Doppler formula (I. 14). This red shift of the spectral lines, the so-called 'transverse' Doppler effect to which we shall come back later (§ 25, p. 62), is a direct consequence of the retardation of moving clocks described by (36), and any experiment which permits an experimental verification of this effect is therefore simultaneously an experimental proof of formula (36).

21. Transformation of particle velocities

Let us again consider the two systems of inertia S and S' (Fig. 8) whose space-time coordinates are connected by (24) and (24′). The motion of an arbitrarily moving particle will then in S be described by a set of equations

$$x = x(t), \quad y = y(t), \quad z = z(t). \tag{41}$$

In S' the same motion is described by functions

$$x' = x'(t'), \quad y' = y'(t'), \quad z' = z'(t'), \tag{41'}$$

which may be obtained from the functions (41) by means of the Lorentz transformations (24) The momentary velocity of the particle relative to S' is defined by

$$\left.\begin{aligned} \mathbf{u}' &= (u'_x, u'_y, u'_z) = \left(\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'}\right) \\ u' &= (u'^2_x + u'^2_y + u'^2_z)^{\frac{1}{2}} \end{aligned}\right\}. \tag{42}$$

Analogously, the corresponding quantities in S are given by

$$\left.\begin{aligned} \mathbf{u} &= (u_x, u_y, u_z) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \\ u &= (u^2_x + u^2_y + u^2_z)^{\frac{1}{2}} \end{aligned}\right\}. \tag{43}$$

Differentiation of the special Lorentz transformation (24) now gives

$$\left.\begin{aligned} dx' &= \gamma(dx - v\,dt), \qquad dy' = dy, \qquad dz' = dz \\ dt' &= \gamma(dt - v\,dx/c^2) \end{aligned}\right\}, \tag{44}$$

from which we immediately obtain

$$u'_x = \frac{u_x - v}{1 - vu_x/c^2}, \qquad u'_y = \frac{u_y(1-v^2/c^2)^{\frac{1}{2}}}{1 - vu_x/c^2}, \qquad u'_z = \frac{u_z(1-v^2/c^2)^{\frac{1}{2}}}{1 - vu_x/c^2}. \tag{45}$$

These equations for the transformation of the velocity reduce to the ordinary transformation equations (I. 4) in the limit as $c \to \infty$. The relations inverse to (45) are obtained in the usual way by interchanging the primed and the unprimed variables and substituting $-v$ for v.

If we choose the z-axis so that $\mathbf{u}$ (and thus also $\mathbf{u}'$) is perpendicular to the z-axis, (45) can be written in the form

$$u'\cos\vartheta' = \frac{u\cos\vartheta - v}{1-uv\cos\vartheta/c^2}, \qquad u'\sin\vartheta' = \frac{u\sin\vartheta(1-v^2/c^2)^{\frac{1}{2}}}{1-uv\cos\vartheta/c^2},$$

where ϑ and ϑ' are the angles between the x-axes and the vectors $\mathbf{u}$ and $\mathbf{u}'$, respectively.

From these equations we obtain directly

$$\tan\vartheta' = \frac{\sin\vartheta(1-v^2/c^2)^{\frac{1}{2}}}{\cos\vartheta - v/u}, \tag{46}$$

$$u' = u\frac{\{1-2v\cos\vartheta/u+v^2/u^2-v^2\sin^2\vartheta/c^2\}^{\frac{1}{2}}}{1-uv\cos\vartheta/c^2}. \tag{47}$$

These equations are the relativistic generalization of the equations (I. 5) and (I. 6). From (47) we obtain by an elementary calculation the formula

$$(1-vu_x/c^2)(1-u'^2/c^2)^{\frac{1}{2}} = (1-u^2/c^2)^{\frac{1}{2}}(1-v^2/c^2)^{\frac{1}{2}}. \tag{48}$$

If $\mathbf{u}$, and thus also $\mathbf{u}'$, are parallel to the x-axis, we get from (45) the relativistic addition theorem for velocities

$$u' = \frac{u-v}{1-uv/c^2}, \qquad u = \frac{u'+v}{1+vu'/c^2}. \tag{49}$$

For $u' = c$ it gives also $u = c$.

From the Lorentz transformations it follows directly that no systems of inertia S' can exist for which $v > c$, since the equations (24) as well as the expressions for the Lorentz contraction and the retardation of clocks would become imaginary in this case. But it can be shown, furthermore, that particles (or, more generally, signals) cannot move with a velocity greater than c relative to any inertial system, since this would lead to absurd results. Let us assume for a moment that we were able to emit signals with a velocity greater than the velocity of light. At the time $t = t' = 0$, where the two systems of coordinates S and S' in Fig. 8 coincide, we could then send a signal from the common origins O, O' along the negative x'-axis with a constant velocity $u' > c$ relative to S'. At the time $t'_1 > 0$, this signal would arrive at a point P on the negative x'-axis with the coordinate $x'_p = -u't'_1$. The space-time coordinates of this event in S are, according to (24'),

$$\left.\begin{aligned} x_p &= \gamma(x'_p+vt'_1) = -\gamma t'_1(u'-v) < 0 \\ t_1 &= \gamma(t'_1+vx'_p/c^2) = \gamma t'_1(1-vu'/c^2) \end{aligned}\right\}. \tag{50}$$

Immediately after its arrival in P the signal is supposed to be sent back to O with a velocity $w > c$ relative to S. The motion of the signal is thus described by the equation

$$x = w(t-t_1)+x_p. \tag{51}$$

This signal will arrive at the origin O of S at a time t_2 which is obtained from this equation by putting $x = 0$, thus at the time

$$t_2 = t_1-x_p/w = \gamma t_1'\{1-u'v/c^2+(u'-v)/w\}. \tag{52}$$

If we now choose u' and w so that

$$u' > c^2/v, \qquad w > \frac{u'-v}{u'v/c^2-1} \tag{53}$$

we could obtain that
$$t_2 < 0, \tag{54}$$
i.e. that, at the return of the signal to O, the clock at O records a number which is smaller than that recorded by the same clock at the moment of departure of the signal. Obviously this is impossible, and therefore we can infer that in nature no signals can exist which move with a velocity greater than the velocity of light relative to any system of inertia. This represents a general statement regarding the fundamental laws of nature. In particular, as we shall see later (Chap. III, § 29, p. 76), a material particle according to relativistic mechanics can never reach velocities larger than c.

By differentiation of (25) and (25′) we obtain the following transformation equations for velocities in the case of Lorentz transformations without rotation

$$\mathbf{u}' = \frac{(1-v^2/c^2)^{\frac{1}{2}}\mathbf{u}+\{[1-\sqrt{(1-v^2/c^2)}](\mathbf{u}\,.\,\mathbf{v})/v^2-1\}\mathbf{v}}{1-(\mathbf{v}\,.\,\mathbf{u})/c^2}, \tag{55}$$

$$\mathbf{u} = \frac{(1-v^2/c^2)^{\frac{1}{2}}\mathbf{u}'+\{[1-\sqrt{(1-v^2/c^2)}](\mathbf{u}'\,.\,\mathbf{v})/v^2+1\}\mathbf{v}}{1+(\mathbf{v}\,.\,\mathbf{u}')/c^2}. \tag{55'}$$

These equations are reduced to (I. 3) in the limit as $c \to \infty$.

From (55) we deduce

$$(1-\mathbf{v}\,.\,\mathbf{u}/c^2)(1-u'^2/c^2)^{\frac{1}{2}} = (1-v^2/c^2)^{\frac{1}{2}}(1-u^2/c^2)^{\frac{1}{2}} \tag{56}$$

in agreement with (48).

When $\mathbf{u}'$ is perpendicular to $\mathbf{v}$, i.e. for $\mathbf{u}'.\mathbf{v} = 0$, (55′) reduces to

$$\mathbf{u} = \mathbf{v}+(1-v^2/c^2)^{\frac{1}{2}}\mathbf{u}' \tag{57}$$

and when $\mathbf{u}$ is parallel to $\mathbf{v}$ we get back to the formula (49).

22. Successive Lorentz transformations. The Thomas precession

Let us now consider three inertial systems S, S', and S'' of which S' moves with the velocity $\mathbf{v}$ relative to S, while S'' moves with the velocity $\mathbf{u}'$ relative to S'. The connexion between the coordinates (x, t) in S and (x', t') in S' is then given by a (generally inhomogeneous) Lorentz transformation. In the same way, the connexion between (x', t') and the coordinates (x'', t'') in S'' is represented by a Lorentz transformation. By elimination of the four variables (x', t') between these eight equations we obtain the connexion between (x, t) and (x'', t'') and, for physical reasons, this relation must also be a Lorentz transformation. Mathematically this is expressed by the statement that the Lorentz transformations form a *group*. If the origins in S and S' coincide at the time $t = t' = 0$, and if the same is the case with S' and S'' at the time $t' = t'' = 0$, the origins in S and S'' will of course coincide at the time $t = t'' = 0$. This shows that the *homogeneous* Lorentz transformations form a subgroup. It is also clear that the pure spatial rotations of the Cartesian axes without any change of the system of reference form a subgroup.

In non-relativistic kinematics the Galilean transformations without rotation of the Cartesian axes also form a subgroup. This is, however, not the case in relativistic kinematics, for if we combine two Lorentz transformations without rotation the resultant Lorentz transformation will in general correspond to a change of orientation of the Cartesian axes. Let the transition from S to S' be given by the Lorentz transformation (25), while the transition from S' to S'' is described by the equation obtained from (25) after replacing $(\mathbf{x}, t, \mathbf{v})$ by $(\mathbf{x}', t', \mathbf{u}')$ and $(\mathbf{x}', t')$ by $(\mathbf{x}'', t'')$. By elimination of $\mathbf{x}'$, t' we then obtain a Lorentz transformation of the type of (28 *b*), viz.

$$\mathbf{x}'' = \mathfrak{D}^{-1}\mathbf{x} - \mathbf{w}''\left\{\frac{\mathbf{x}\cdot\mathbf{w}}{w^2}\left[\frac{1}{\sqrt{(1-w^2/c^2)}} - 1\right] - \frac{t}{\sqrt{(1-w^2/c^2)}}\right\}, \tag{58}$$

where the operator $\mathfrak{D}$ in general is different from the unity operator. $\mathbf{w}$ is the velocity of the system S'' relative to S and $\mathbf{w}''$ is the velocity of S relative to S''. Since the transformations from S to S' and from S' to S'' were Lorentz transformations without rotation, the velocity of S' relative to S'' is equal to $-\mathbf{u}'$ while the velocity of S relative to S' is equal to $-\mathbf{v}$. We now obtain the velocity $\mathbf{w}$ of S'' relative to S from (55'), identifying u with w in this formula, i.e.

$$\mathbf{w} = \frac{(1-v^2/c^2)^{\frac{1}{2}}\mathbf{u}' + \mathbf{v}\{(\mathbf{u}'\cdot\mathbf{v}/v^2)[1-(1-v^2/c^2)^{\frac{1}{2}}]+1\}}{1+(\mathbf{u}'\cdot\mathbf{v})/c^2}. \tag{59}$$

From the same equation we get the velocity $\mathbf{w}''$ of S relative to S'' by replacing $\mathbf{v}$ and $\mathbf{u}'$ by $-\mathbf{u}'$ and $-\mathbf{v}$, respectively; thus we have

$$\mathbf{w}'' = -\frac{(1-u'^2/c^2)^{\frac{1}{2}}\mathbf{v}+\mathbf{u}'\{(\mathbf{u}'.\mathbf{v}/u'^2)[1-(1-u'^2/c^2)^{\frac{1}{2}}]+1\}}{1+(\mathbf{u}'.\mathbf{v})/c^2}. \tag{59'}$$

According to (29) we have in our case

$$\mathfrak{D}\mathbf{w}'' = -\mathbf{w}, \tag{60}$$

and a comparison of (59) and (59′) shows that the rotation operator $\mathfrak{D}$ in general is different from the unity operator. Only if $\mathbf{u}'$ is parallel to $\mathbf{v}$, say equal to $k\mathbf{v}$, we get from (59) and (59′)

$$\mathbf{w}'' = -\mathbf{w} = -\frac{1+k}{1+kv^2/c^2}\mathbf{v},$$

i.e. in this case $\mathfrak{D} = 1$ and the combined Lorentz transformation is also without rotation.

Let us now consider the case where the transition from S' to S'' is an infinitesimal transformation, i.e. where u' is infinitesimal. Neglecting all terms of higher than first order in $\mathbf{u}'$, the transformation from S' to S'' is reduced to

$$\mathbf{x}'' = \mathbf{x}'-\mathbf{u}'t', \qquad t'' = t'-(\mathbf{u}'.\mathbf{x}')/c^2 \tag{61}$$

and for $\mathbf{w}$ and $\mathbf{w}''$ we obtain from (59) and (59′)

$$\left.\begin{aligned} \mathbf{w} &= \mathbf{v}+\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}\left\{\mathbf{u}'+\mathbf{v}\frac{(\mathbf{v}.\mathbf{u}')}{v^2}\left[\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}-1\right]\right\} \\ \mathbf{w}'' &= -\left(\mathbf{v}+\mathbf{u}'-\mathbf{v}\frac{(\mathbf{v}.\mathbf{u}')}{c^2}\right) \\ w^2 &= w''^2 = v^2+2(\mathbf{v}.\mathbf{u}')\left(1-\frac{v^2}{c^2}\right) \end{aligned}\right\}. \tag{62}$$

By substituting the expressions (25) for x' and t' in (61) we obtain an equation which after a simple, but lengthy, calculation may be written in the form (58) with

$$\left.\begin{aligned} \mathfrak{D}^{-1}\mathbf{x} &= \mathbf{x}+\frac{1}{v^2}\{(1-v^2/c^2)^{-\frac{1}{2}}-1\}\{(\mathbf{v}\times d\mathbf{v})\times\mathbf{x}\} \\ d\mathbf{v} &= \mathbf{w}-\mathbf{v} \end{aligned}\right\}. \tag{63}$$

Thus, we have

$$\left.\begin{aligned} \mathfrak{D}\mathbf{x} &= \mathbf{x}+(\mathbf{\Omega}\times\mathbf{x}) \\ \mathbf{\Omega} &= -\frac{1}{v^2}\{(1-v^2/c^2)^{-\frac{1}{2}}-1\}(\mathbf{v}\times d\mathbf{v}) \end{aligned}\right\}. \tag{64}$$

The rotation operator $\mathfrak{D}$ thus represents an infinitesimal rotation around the direction of the vector $\mathbf{\Omega}$, the angle of rotation being equal to the magnitude $|\mathbf{\Omega}|$ of the vector $\mathbf{\Omega}$. With $\mathfrak{D}$ given by (64) and $\mathbf{w}$ and $\mathbf{w}''$

given by (62) it is easily verified that the equation (60) is satisfied to the approximation considered.

Let us now consider a point compass, i e. a material particle which in some way or other defines a direction. It may be assumed that a classical electron with spin represents such a point compass. If the velocity of the particle relative to S is $\mathbf{v} = \mathbf{v}(t)$, and if we put $d\mathbf{v} = \dot{\mathbf{v}}(t)\,dt$ in (64), the systems S' and S'' in the above-mentioned considerations will be momentary rest systems of inertia for the particle at the times t and $t+dt$, respectively. Since the transition from S' to S'' represents an infinitesimal Lorentz transformation without rotation, it is natural to assume that the direction shown by the compass at the time $t+dt$ has the same orientation relative to the axes in S'' as it has at the time t relative to the axes in S', provided that the forces on the particle do not exert any torque on the compass.

Now, if we put $d\mathbf{v} = \dot{\mathbf{v}}\,dt$, the rotation vector $\mathbf{\Omega}$ defined by (64) represents the rotation which has to be applied to the axes in S at the time $t+dt$ in order to give them the same orientation as have the axes in S''. Since, furthermore, the direction of the compass relative to the rest system is constant, this means that the direction of the compass relative to S is turned through an angle corresponding to the rotation vector $\mathbf{\Omega}$. In other words, the compass performs a precession relative to S with the velocity of precession

$$\boldsymbol{\omega} = -\frac{1}{v^2}\{(1-v^2/c^2)^{-\frac{1}{2}}-1\}(\mathbf{v}\times\dot{\mathbf{v}}), \tag{65}$$

where $\dot{\mathbf{v}} = d\mathbf{v}/dt$ is the acceleration of the point compass. When $v \ll c$ we obtain to a first approximation

$$\boldsymbol{\omega} = -\frac{\mathbf{v}\times\dot{\mathbf{v}}}{2c^2}. \tag{66}$$

This precession phenomenon was studied for the first time by Thomas† and is therefore called the Thomas precession.

23. Transformation of the characteristics of a wave according to the theory of relativity

Let us again consider a plane wave with the wave normal $\mathbf{n}$ in the xy-plane of a system of coordinates S and with the frequency ν and the velocity w relative to S. It is described by one or several wave functions of the form

$$\Psi = A\cos 2\pi\nu\left(t-\frac{x\cos\alpha+y\sin\alpha}{w}\right), \tag{67}$$

† L. W. Thomas, *Phil. Mag.* (7), **3**, 1 (1927).

where α is the angle between the wave normal $\mathbf{n}$ and the x-axis. In a system of coordinates S' moving in the direction of the x-axis with the velocity v relative to S (cf. Fig. 1), the wave is described by corresponding functions obtained from (67) when replacing the quantities occurring in the phase by the corresponding quantities measured in the system of coordinates S'. From exactly the same argument as that used in § 3 it follows that the phase must be an invariant, i.e. the equation

$$\nu\left(t-\frac{x\cos\alpha+y\sin\alpha}{w}\right)=\nu'\left(t'-\frac{x'\cos\alpha'+y'\sin\alpha'}{w'}\right) \tag{68}$$

must be valid for all points in space and at all times. Now using the connexion between the space-time coordinates in S and S' given by the Lorentz transformations (24') we can eliminate the variables x, y, t in this equation and get

$$\frac{1-v\cos\alpha/w}{(1-v^2/c^2)^{\frac{1}{2}}}\nu t'-\frac{\cos\alpha-vw/c^2}{(1-v^2/c^2)^{\frac{1}{2}}}\frac{\nu x'}{w}-\frac{\nu\sin\alpha}{w}y'$$
$$=\nu' t'-\frac{\nu'\cos\alpha'}{w'}x'-\frac{\nu'\sin\alpha'}{w'}y', \tag{69}$$

an equation which must be satisfied for all values of the independent variables t', x', y'. This is possible only if the coefficients of t', x', y', respectively, on both sides are equal. Therefore we must have

$$\left.\begin{aligned}\nu'=\nu\frac{1-v\cos\alpha/w}{(1-v^2/c^2)^{\frac{1}{2}}}&=\nu\frac{1-(\mathbf{v}\cdot\mathbf{n})/w}{(1-v^2/c^2)^{\frac{1}{2}}}\\ \frac{\nu'\sin\alpha'}{w'}&=\frac{\nu\sin\alpha}{w}\\ \frac{\nu'\cos\alpha'}{w'}&=\frac{\nu(\cos\alpha-vw/c^2)}{w(1-v^2/c^2)^{\frac{1}{2}}}\end{aligned}\right\}. \tag{70}$$

From these equations we obtain

$$\tan\alpha'=\frac{\sin\alpha(1-v^2/c^2)^{\frac{1}{2}}}{\cos\alpha-vw/c^2}, \tag{71}$$

$$w'=\frac{w-v\cos\alpha}{\sqrt{\left(1-\frac{2vw\cos\alpha}{c^2}+\frac{v^2w^2}{c^4}-\frac{v^2\sin^2\alpha}{c^2}\right)}}. \tag{72}$$

The inverse relations are obtained in the usual way by interchanging the primed and the unprimed quantities and replacing v by $-v$. In the limit as $c\to\infty$, the equations (70), (71), and (72) reduce to the non-relativistic formulae (I. 14), (I. 16), and (I. 19).

A comparison of the transformation equations (46) and (47) for the velocity and direction of a material particle with the formulae (71) and (72) shows that (46) and (47) become (71) and (72), respectively, when we put $u = c^2/w$, $u' = c^2/w'$. In other words, the velocity u of a particle and its direction **n** are transformed in the same manner as the corresponding quantities for a wave with the phase velocity $w = c^2/u$ and direction **n**. In his wave theory of elementary particles de Broglie† made use of this circumstance by attributing to a particle with the velocity u and the direction of motion **n** a plane wave with the same direction of propagation **n** and the phase velocity $w = c^2/u$, a procedure which thus is relativistically invariant. If the particle velocity $u = c$ the corresponding wave velocity is $w = c$, which shows that direction and velocity of such a particle are transformed in the same way as are direction and velocity of a plane light wave in vacuum, but this only holds for this particular value of the velocity.

24. The ray velocity in moving bodies

Consider a homogeneous isotropic medium, with the refractive index n, at rest in the system of coordinates S' (Fig. 8). Relative to S the medium moves with the velocity v. In the rest system S', Maxwell's phenomenological equations of electrodynamics in dielectric bodies are valid, and, according to the principle of relativity, this must be true for any constant velocity of S' relative to the fixed stars. The phase velocity of light relative to S' is, therefore, $w' = c/n$ in all directions. In this system of reference the ray velocity must, however, be equal to the phase velocity, since the elementary waves which, according to Huyghens's principle, determine the ray velocity are spherical waves with the constant velocity

$$w' = c/n. \tag{73}$$

This is not so, however, in the system of coordinates S. Let us consider, for instance, the propagation of an elementary wave which, at the time $t = t' = 0$, is emitted from the common origin O, O' of S and S'. In S' the propagation of the wave is described by the equation

$$x'^2+y'^2+z'^2-w'^2t'^2 = 0, \tag{74}$$

where w' is given by (73). By means of the Lorentz transformation (24) we obtain from (74) the following equation for the propagation of the elementary wave in S:

$$\left.\begin{aligned} &(x-at)^2/b+y^2+z^2-bw'^2t^2 = 0 \\ a = v\frac{1-w'^2/c^2}{1-v^2w'^2/c^4},& \qquad b = \frac{1-v^2/c^2}{1-v^2w'^2/c^4} \end{aligned}\right\}. \tag{75}$$

† L. de Broglie, Thèse, Paris 1924.

An elementary wave which is emitted at the time t from a point (x_0, y_0) in the xy-plane of the system S will thus, at the time $t+\Delta t$, have a curve of intersection with the xy-plane which is described by the equation

$$f(x, y;\, x_0, y_0) = (x-x_0-a\,\Delta t)^2/b+(y-y_0)^2-bw'^2\,\Delta t^2 = 0. \qquad (76)$$

We shall here only deal with the case $n > 1$, i e. $w' < c$. In this case $c > a > 0$ and $1 > b > 0$ and the curve of intersection consequently

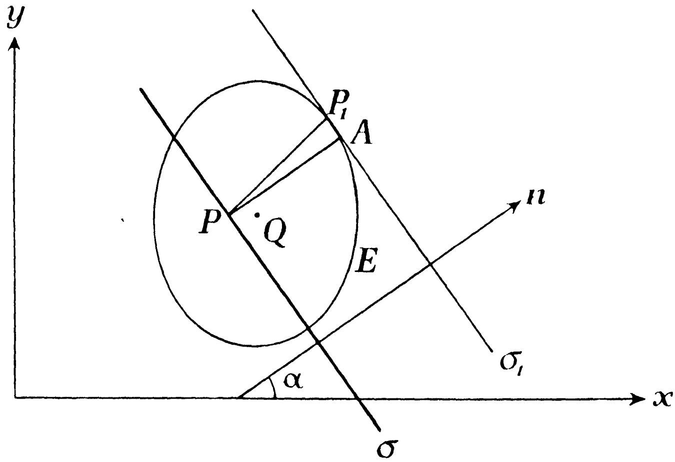

FIG. 9.

is an ellipse with its centre at the point $(x_0+a\,\Delta t, y_0)$ and with the semi-axes $bw'\Delta t$ and $b^{\frac{1}{2}}w'\Delta t$, respectively. The elementary waves are thus dragged along by the medium with the velocity a, and simultaneously the waves are contracted in the direction of motion in the proportion $b^{\frac{1}{2}}$.

Now let us again consider a plane wave with a normal **n** which lies in the xy-plane and makes an angle α with the x-axis. The connexion between α and the direction of the normal in S' is then given by (71) or by the inverse equation

$$\tan\alpha = \frac{(1-v^2/c^2)^{\frac{1}{2}}\sin\alpha'}{\cos\alpha'+vw'/c^2}. \qquad (77)$$

Let σ in Fig. 9 be a wave plane in S whose line of intersection with the xy-plane is given by the equation

$$x_0\cos\alpha+y_0\sin\alpha = C = \text{constant}. \qquad (78)$$

The ellipse E with its centre at the point Q: $(x_0+a\,\Delta t, y_0)$ represents the elementary wave at the time $t+\Delta t$, which is emitted from the point P: (x_0, y_0) at the time t. It is given by the equation (76). The wave plane at the time $t+\Delta t$ is represented by the line σ_1, which is obtained as the envelope of the manifold of ellipses (76) which are obtained by varying the parameters (x_0, y_0) in accordance with (78).

The direction and magnitude of the ray velocity are thus given by the line $\overrightarrow{PP_1}$, where P_1 with coordinates (x_1, y_1) is the point at which E touches the envelope σ_1. We have obviously

$$\overrightarrow{PP_1} = \mathbf{u}\,\Delta t, \tag{79}$$

where $\mathbf{u}$ is the ray velocity. Since P_1 is the limiting point of intersection of two adjacent ellipses of the manifold of curves (76), we see that the coordinates of P_1 besides satisfying the equation (76) must satisfy the equation obtained by varying x_0, y_0 in (76) in accordance with (78), i.e. the equation

$$\frac{\partial f}{\partial x_0}\sin\alpha - \frac{\partial f}{\partial y_0}\cos\alpha = 0$$

or

$$(x-x_0-a\,\Delta t)\sin\alpha = b(y-y_0)\cos\alpha. \tag{80}$$

From the three equations (76), (78), and (80) we obtain by elimination of x_0 and y_0 the equation for the envelope σ_1 in the form

$$x\cos\alpha+y\sin\alpha = C+[a+w'\{b(b+\tan^2\alpha)\}^{\frac{1}{2}}]\cos\alpha\,\Delta t. \tag{81}$$

The last term in (81) gives the distance PA between the two wave planes which must be equal to $w\,\Delta t$, where w is the phase velocity. Therefore we obtain

$$w = \{a+w'[b(b+\tan^2\alpha)]^{\frac{1}{2}}\}\cos\alpha. \tag{82}$$

By inserting the expressions (75) and (77) for a, b, and α it is easily verified that (82) is identical with the inverse of the equation (72).

According to (79) we have

$$u_x = \frac{x_1-x_0}{\Delta t}, \qquad u_y = \frac{y_1-y_0}{\Delta t}, \tag{83}$$

and, since $x = x_1$, $y = y_1$ is a solution of the equations (76) and (80), we obtain by a simple calculation

$$u_x = a+\frac{b^{\frac{1}{2}}w'}{(b+\tan^2\alpha)^{\frac{1}{2}}}, \qquad u_y = \frac{b^{\frac{1}{2}}\tan\alpha\, w'}{(b+\tan^2\alpha)^{\frac{1}{2}}}. \tag{84}$$

When inserting the expressions (75) and (77) for a, b, α and remembering that the ray velocity $\mathbf{u}'$ is equal to the phase velocity $\mathbf{w}'$ in S', we find that (84) is reduced to

$$u_x = \frac{u'_x+v}{1+u'_x v/c^2}, \qquad u_y = \frac{(1-v^2/c^2)^{\frac{1}{2}}u'_y}{1+u'_x v/c^2}. \tag{85}$$

The equations (85) are identical with the relations inverse to (45), which shows that the ray velocity (just as in the absolute ether theory, cf. (I. 33)) is transformed in the same way as the velocity of a material particle. When ϑ and $\vartheta' = \alpha'$ denote the angles between the x-axes and the direction of the ray measured in S and S', respectively, we have therefore also for the ray velocities the equations (46) and (47) connecting the quantities u, u', ϑ, and ϑ'. By solving (47) with respect to u, we obtain

$$u = \frac{u'\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}\left\{1-\frac{v^2}{u'^2}+\frac{(\mathbf{v}.\mathbf{e})^2}{u'^2}\left(1-\frac{u'^2}{c^2}\right)\right\}^{\frac{1}{2}}+(\mathbf{v}.\mathbf{e})\left(1-\frac{u'^2}{c^2}\right)}{1-\frac{v^2}{c^2}+\frac{(\mathbf{v}.\mathbf{e})^2}{c^2}\left(1-\frac{u'^2}{c^2}\right)}. \tag{86}$$

Here

$$u' = c/n \tag{87}$$

and $\mathbf{e}$ is a unit vector in the direction of the ray, i.e.

$$\mathbf{v}.\mathbf{e} = v\cos\vartheta. \tag{88}$$

In vacuum, i.e. for $n = 1$, $u' = w' = c$, u and w are consequently equal to c, and the formulae (46) and (47) for the ray velocity become identical with the equations (71) and (72) for the phase velocity. According to the theory of relativity, the ray velocity in vacuum is thus identical with the phase velocity in *any* inertial system. In the absolute ether theory this was the case only in the absolute system. This difference is due to the fact that the elementary waves in vacuum according to the theory of relativity are spherical waves with a fixed centre in any system of inertia. (For $w' = c$ we get from (75) $a = 0$ and $b = 1$.) We shall see later (Chapters V, VII) that the ray velocity is equal to the velocity with which the electromagnetic energy is flowing in an electromagnetic wave. The energy current density is, however, given by Poynting's vector, and for a plane wave in vacuum it is found that Poynting's vector lies in the direction of the wave normal in all systems of inertia.

This identity of ray velocity and phase velocity applies only to the vacuum. In a refractive medium we have in general to distinguish between the two velocities; only when the ray is parallel to the direction of motion of the medium, both the ray velocity and the phase velocity are, according to (85) or (86) and (72), given by the same formula, viz.

$$u = w = \frac{c/n \pm v}{1 \pm v/cn}, \tag{89}$$

where the plus or the minus signs should be taken depending on whether the ray travels in the same direction as the medium or opposite to it.

25. The Doppler effect, the aberration of light, and the dragging phenomenon according to the theory of relativity

The relativistic formula for the Doppler effect is obtained from (70) by identifying the inertial systems S and S' with the rest systems of the observer and of the light source, respectively. The frequency ν is then the frequency measured by the observer, and ν' is equal to the proper frequency ν^0 of the light source. Now, in vacuum, $w = c$ and the direction of the ray is equal to the direction of the wave normal. Thus, we obtain

$$\nu^0 = \nu \frac{1-(\mathbf{v}.\mathbf{e})/c}{\sqrt{(1-v^2/c^2)}}, \tag{90}$$

where $\mathbf{v}$ is the velocity of the light source relative to the observer and $\mathbf{e} = \mathbf{n}$ gives the direction of the light in the system of the observer. For $(\mathbf{v}.\mathbf{e}) = 0$, i e. when the direction of the observed light is perpendicular to the direction of motion of the light source, we get the formula for the transverse Doppler effect $\nu = \nu^0(1-v^2/c^2)^{\frac{1}{2}}$, which, as mentioned in § 20, is a direct expression for the retardation of moving clocks. As mentioned in § 5, the formula (90) has been experimentally verified with high accuracy by Ives,† who measured the frequency of the light emitted from rapidly moving ions.

The formula (46) connecting the directions of a light ray in two inertial systems S and S' leads immediately to the relativistic aberration formula, when we take S and S' to denote the systems of coordinates in which the fixed stars and the earth, respectively, are at rest. At a point P' outside the atmosphere of the earth we have $u = c$ and, when introducing the angles $\theta = \vartheta-\pi$ and $\theta' = \vartheta'-\pi$ between the direction of motion of the earth and the actual and apparent directions to the star, we obtain the relativistic aberration formula

$$\tan\theta' = \frac{\sin\theta(1-v^2/c^2)^{\frac{1}{2}}}{\cos\theta+v/c}. \tag{91}$$

It deviates from the equation (I. 61) of the ether theory by quantities of second order only, a deviation which is negligible compared with the present accuracy of the measurements.

Since the atmosphere of the earth is at rest relative to S', it follows at once from the principle of relativity that the ray during its travel from the point P' to the telescope does not undergo any further aberration. According to the theory of relativity this should even hold exactly, while Lorentz's absolute theory of electrons gave this result to a first approximation only (cf. § 10).

† See ref., p. 10.

Another essential difference in the relativistic treatment is that here the aberration also appears in the direction of the phase velocity. While, in the absolute theory, the wave normal according to (I. 28) has the same direction in S and in S', the relativistic equation (71) for the transformation of the wave normal *in vacuo* is identical with the transformation equation for the direction of the ray (46).

Neglecting all terms of higher than first order in v we obtain from equation (86) for the ray velocity in a medium moving with constant velocity $\mathbf{v}$ relative to an arbitrary inertial system S the simple formula

$$u = u'+\mathbf{v}.\mathbf{e}(1-u'^2/c^2) = c/n+\alpha(\mathbf{v}.\mathbf{e}), \tag{92}$$

where $\alpha = 1-1/n^2$ is Fresnel's dragging coefficient. Equation (92) is, in this approximation, identical with Fresnel's equation (I. 48) for the velocity of light in the 'absolute' system of coordinates. This formula thus follows directly from the principle of relativity without any hypotheses regarding the atomic structure of the medium. It should be stressed, however, that Fresnel's theory is true to a first approximation only in the 'absolute' system. In any other system Fresnel's equations are not valid even to a first approximation. According to the theory of relativity we have, for example, $u' = c/n$ in the rest system of the medium whereas u' according to Fresnel's theory is given by the more complicated formula (I. 46).

Equation (92) has been verified not only by Fizeau's experiments discussed in § 8, but also with high accuracy by Zeeman.† In these latter experiments the velocity of light in a rapidly moving quartz rod was measured. Indeed Zeeman's measurements were so accurate that it was necessary to take into account the effect which, according to Lorentz's electron theory, should occur in a dispersive medium (cf. § 9). Since $n = n(\nu)$ in such substances is dependent on ν, and since ν on account of the Doppler effect depends on the system of coordinates, it must be specified which value for ν, and thus for n, should be introduced into equation (92). It follows from the derivation of (86) and (92) given above that we have in (92) to use the value of n corresponding to the frequency ν' in the rest system of the medium. When we neglect all quantities of order higher than the first, the connexion between ν' and ν, according to (70), is given by

$$\nu'-\nu = -\nu n(\mathbf{v}.\mathbf{e})/c, \tag{93}$$

the difference between the directions of the ray and of the wave normal

† P. Zeeman, *Amst. Versl.* **23**, 245 (1914); **24**, 18 (1915).

being of second order in v. Hence,

$$\frac{c}{n(\nu')} = \frac{c}{n(\nu)} + \frac{c}{n^2}\frac{dn}{d\nu}\frac{n\nu(\mathbf{v}.\mathbf{e})}{c},$$

where $n = n(\nu)$ means the value of n for the frequency ν. By using this formula in (92) we obtain

$$u = \frac{c}{n} + (\mathbf{v}.\mathbf{e})\left(1 - \frac{1}{n^2} + \frac{\nu}{n}\frac{dn}{d\nu}\right). \tag{94}$$

This formula was in complete agreement with Zeeman's measurements.

The reflection of light by moving mirrors and the refraction of light at the transition between moving media can be treated in a similar way on the basis of the theory of relativity. In a system of inertia S', in which the mirror or the medium is at rest, the usual law for reflection of a light ray by a mirror and the well-known law of refraction at the transition between two media are valid. The corresponding laws in the system S are then directly obtained by applying the transformation equations (46) and (47) for the ray velocity. Experiments on reflection by moving mirrors have been performed by Sagnac† and others. The result of these experiments is in accordance with the theory.

In all these experiments, only effects of first order could be measured, since the velocities with which the mirrors and the refractive media could be moved were so small compared with the velocity of light that the terms of second order in equation (86) were negligible compared with the accuracy of the measurements. In this connexion it is instructive to consider a slightly generalized Michelson experiment in which the whole apparatus is filled with a refractive medium. From the principle of relativity it is clear that also in this case no displacement of the interference fringes will occur when the apparatus is rotated (cf. Fig. 6, p. 27). This result is also obtained when we calculate the times t_1' and t_2' which the two rays take to traverse the paths PS_1P and PS_2P in the system of inertia S' in which the apparatus is at rest; for in this system the velocity of light is the same, viz. c/n in all directions and therefore $t_1' = t_2'$.

Now it must be required, of course, that the same zero result is obtained when the effect is treated from the point of view of an observer in a system S, relative to which the whole apparatus has a velocity $\mathbf{v}$. The corresponding time-intervals t_1 and t_2, measured by clocks in S, may now be obtained from the formula (86) and it is easily seen that again $t_1 = t_2$. The time which the ray 2 needs to travel from P to S_2 is obviously $\frac{1}{2}t_2$. During this time the apparatus has moved over

† G. Sagnac, *C.R.* **157**, 708, 1410 (1913), *Journ. de Phys.* (5), **4**, 177 (1914).

a distance $\frac{1}{2}vt_2$ in S, and the ray has traversed a length $\frac{1}{2}ut_2$, where the ray velocity u is determined by (86) or (47). In Fig. 10 these distances are shown as PP^* and PS_2^*, P^* and S_2^* denoting the positions occupied by P and S_2 after the time $\frac{1}{2}t_2$. From the figure we see that the angle ϑ between the ray and the direction of $\mathbf{v}$ is given by

$$\cos\vartheta = v/u.$$

Substituting this in (47) we obtain, by solving this equation with respect to u^2,

$$u^2 = u'^2\left(1-\frac{v^2}{c^2}\right)+v^2. \tag{95}$$

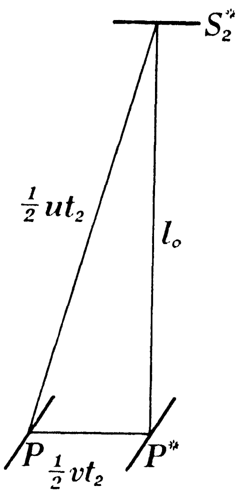

FIG. 10.

By means of the Pythagorean theorem applied to the triangle $PP^*S_2^*$ we obtain, moreover, from Fig. 10

$$\tfrac{1}{2}t_2 = \frac{l_0}{(u^2-v^2)^{\frac{1}{2}}}, \tag{96}$$

where l_0 is the distance between the plate P and the mirror S_2. From (96) and (95) we then get for the time t_2 which the ray 2 needs to travel from P to S_2 and back to P

$$t_2 = \frac{2l_0}{u'(1-v^2/c^2)^{\frac{1}{2}}}. \tag{97}$$

Proceeding to the calculation of the time which the ray 1 needs to travel from P to S_1 and back to P, we shall here use for the velocity u the expressions (89), since the ray 1 all the time moves parallel to the direction of motion of the apparatus. If we denote the time which this ray needs to travel from P to S_1 by t_1^+, the mirror S_1 will have moved a distance vt_1^+ during this time and the total distance which the ray has travelled relative to S is equal to $l+vt_1^+$, where

$$l = l_0(1-v^2/c^2)^{\frac{1}{2}} \tag{98}$$

is the distance between P and S_1 measured in S. Therefore we have

$$\frac{c/n+v}{1+v/cn}t_1^+ = l+vt_1^+$$

or

$$t_1^+ = \frac{l_0 n(1+v/cn)}{c(1-v^2/c^2)^{\frac{1}{2}}}. \tag{99}$$

The time t_1^- which the ray needs to travel from S_1 back to S is obtained simply by replacing v by $-v$, i.e.

$$t_1^- = \frac{l_0 n(1-v/cn)}{c(1-v^2/c^2)^{\frac{1}{2}}}.$$

The total time t_1 is then

$$t_1 = t_1^+ + t_1^- = \frac{2l_0 n}{c(1-v^2/c^2)^{\frac{1}{2}}}. \tag{100}$$

Thus t_1 is equal to t_2, as was to be required.

Even if this result only shows that the theory is consistent, it also allows us to make a choice between Fresnel's formula (I. 47) and the relativistic equation (86) which differ from each other in the terms of second order. If we use (I. 47) instead of (86) in the calculations just performed, we find for the times t_1 and t_2 the expressions

$$t_1 = \frac{2nl}{c(1-v^2/c^2n^2)}, \qquad t_2 = \frac{2nl_0}{c(1-v^2/c^2n^2)^{\frac{1}{2}}}, \tag{101}$$

which shows that for $n \neq 1$ we would obtain $t_2 \neq t_1$ even if we assume that the distance PS_1 is contracted according to the Lorentz contraction formula (98).

In the same way as the negative result of the usual Michelson experiment can be regarded as an experimental verification of Lorentz's formula (98), a negative result of the corresponding experiment, where the apparatus is filled with a strongly refractive medium, would mean a verification of the relativistic formula (86) also as regards terms of higher order. The same considerations can be applied to Hoek's experiment discussed in § 8.

Summarizing, it can be stated that relativistic kinematics in all details is in agreement with the experimental results. It allows a simple explanation of all dragging phenomena without any *ad hoc* hypotheses, and it gives a formula for the Doppler effect which, in contradistinction to the formula of the ether theory, is in accordance with the experimental results.

III

RELATIVISTIC MECHANICS

26. Momentum and mass of a particle

As mentioned in the conclusion of § 18, it is necessary to change Newton's fundamental equations of mechanics in order to bring them into accordance with the principle of relativity. In the domain, where all velocities are small compared with c, relativistic mechanics will, however, go over into Newtonian mechanics. It is therefore natural to assume that such fundamental concepts of Newtonian mechanics as momentum and mass of a material particle also have a meaning in relativistic mechanics.

Therefore, to a material particle moving with the velocity $\mathbf{u}$ relative to a system of inertia S, we shall assign a momentum vector $\mathbf{p}$ proportional to $\mathbf{u}$

$$\mathbf{p} = m\mathbf{u}. \tag{1}$$

The proportionality factor m is called the mass of the particle. To make room for the above-mentioned generalization of mechanics we shall not, however, assume beforehand that m is a constant, but we make the assumption that m is a universal function, $f(u)$, of the magnitude $u = |\mathbf{u}|$ of the velocity vector, thus

$$m = m(u) = f(u). \tag{2}$$

If the velocity of the particle relative to another system of inertia S' is $\mathbf{u}'$, the momentum and mass of the particle relative to S' must be given by

$$\mathbf{p}' = m'\mathbf{u}', \tag{3}$$

where

$$m' = m'(u') = f(u') \tag{4}$$

is the same function of u' as m is of u. This follows from the principle of relativity according to which all systems of inertia have to be treated on the same footing, so that *any relation between physical quantities shall be form-invariant.*

It will now be our primary task to determine the function f. As we shall see, this function is uniquely determined when we require that the theorem of conservation of momentum shall hold in any system of inertia.† Let S and S' be two systems of inertia with the relative velocity $\mathbf{v}$, and consider a collision between two identical particles 1 and 2 which before the collision have the velocities $\mathbf{u}_1$ and $\mathbf{u}_2$ relative to S. The corresponding velocities relative to S' are then determined by (II. 55).

† G. N. Lewis and R. C. Tolman, *Phil. Mag.* **18**, 510 (1909).

Let us now choose the velocities before the collision so that

$$\mathbf{u}_2' = -\mathbf{u}_1. \tag{5}$$

By means of (II. 55) it follows then at once that also

$$\mathbf{u}_1' = -\mathbf{u}_2. \tag{6}$$

After the collision the particles will have other velocities, $\bar{\mathbf{u}}_1$ and $\bar{\mathbf{u}}_2$ relative to S, and $\bar{\mathbf{u}}_1'$ and $\bar{\mathbf{u}}_2'$ relative to S'. We shall in particular consider a collision in which the final and initial velocities of particle 1 have opposite directions, i.e.

$$\bar{\mathbf{u}}_1 = -\alpha \mathbf{u}_1, \tag{7}$$

where α is a positive number. For symmetry reasons we must then also have

$$\bar{\mathbf{u}}_2' = -\alpha \mathbf{u}_2' \tag{8}$$

with the same proportionality factor α as in (7), for, according to (5), the motion of particle 2 relative to an observer in S' must be the same as the motion of particle 1 relative to an observer in S. From (5), (7), and (8) we get

$$\bar{\mathbf{u}}_2' = -\bar{\mathbf{u}}_1, \tag{9}$$

which, according to (II. 55), involves

$$\bar{\mathbf{u}}_1' = -\bar{\mathbf{u}}_2.$$

Assuming that the functional relation (2) is the same before and after the collision, conservation of momentum in S gives

$$f(u_1)\mathbf{u}_1 + f(u_2)\mathbf{u}_2 = f(\bar{u}_1)\bar{\mathbf{u}}_1 + f(\bar{u}_2)\bar{\mathbf{u}}_2. \tag{10}$$

Let us, furthermore, assume that the velocity of particle 1 before the collision is perpendicular to $\mathbf{v}$, i.e.

$$(\mathbf{u}_1 \cdot \mathbf{v}) = 0. \tag{11}$$

According to (5), (7), and (9) this involves

$$(\mathbf{u}_2' \cdot \mathbf{v}) = 0 \tag{12}$$

and

$$(\bar{\mathbf{u}}_1 \cdot \mathbf{v}) = (\bar{\mathbf{u}}_2' \cdot \mathbf{v}) = 0. \tag{13}$$

For the transformation of the velocities $\mathbf{u}_2$ and $\bar{\mathbf{u}}_2$ we can therefore use the simple formula (II. 57). Thus we get

$$\mathbf{u}_2 = \mathbf{u}_2'(1-v^2/c^2)^{\frac{1}{2}} + \mathbf{v} = \mathbf{v} - \mathbf{u}_1(1-v^2/c^2)^{\frac{1}{2}} \tag{14}$$

and

$$u_2^2 = (\mathbf{u}_2 \cdot \mathbf{u}_2) = u_1^2(1-v^2/c^2) + v^2, \tag{15}$$

the cross-terms being zero on account of (11).

In the same way we obtain by means of (9)

$$\left.\begin{aligned} \bar{\mathbf{u}}_2 &= \bar{\mathbf{u}}_2'(1-v^2/c^2)^{\frac{1}{2}} + \mathbf{v} = -\bar{\mathbf{u}}_1(1-v^2/c^2)^{\frac{1}{2}} + \mathbf{v} \\ \bar{u}_2^2 &= \bar{u}_1^2(1-v^2/c^2) + v^2 \end{aligned}\right\}. \tag{16}$$

Introducing (14), (15), and (16) into (10), we get an equation which can be written

$$[f(u_1)-(1-v^2/c^2)^{\frac{1}{2}}f\{\sqrt{[u_1^2(1-v^2/c^2)+v^2]}\}]\mathbf{u}_1+f\{\sqrt{[u_1^2(1-v^2/c^2)+v^2]}\}\mathbf{v}$$
$$= [f(\bar{u}_1)-(1-v^2/c^2)^{\frac{1}{2}}f\{\sqrt{[\bar{u}_1^2(1-v^2/c^2)+v^2]}\}]\bar{\mathbf{u}}_1+f\{\sqrt{[\bar{u}_1^2(1-v^2/c^2)+v^2]}\}\mathbf{v}. \quad (17)$$

Multiplying this vector equation by $\mathbf{v}$ yields, on account of (11) and (13),

$$f\{\sqrt{[u_1^2(1-v^2/c^2)+v^2]}\} = f\{\sqrt{[\bar{u}_1^2(1-v^2/c^2)+v^2]}\}.$$

Hence, assuming that f is a monotonic function of the argument,

$$\bar{u}_1 = u_1, \quad (18)$$

i.e. $\alpha = 1$ in (7). This means, on the other hand, that the terms proportional to $\mathbf{v}$ in (17) cancel each other and that the coefficients of $\bar{\mathbf{u}}_1$ and $\mathbf{u}_1$ are equal. Consequently, equation (17) reduces to

$$[f(u_1)-f\{\sqrt{[u_1^2(1-v^2/c^2)+v^2]}\}(1-v^2/c^2)^{\frac{1}{2}}](\mathbf{u}_1-\bar{\mathbf{u}}_1) = 0. \quad (19)$$

Since, moreover, according to (7) and (18),

$$\mathbf{u}_1-\bar{\mathbf{u}}_1 = 2\mathbf{u}_1 \neq 0,$$

the coefficient of $\mathbf{u}_1-\bar{\mathbf{u}}_1$ in (19) must be zero, i.e.

$$f(u_1) = (1-v^2/c^2)^{\frac{1}{2}}f\{\sqrt{[u_1^2(1-v^2/c^2)+v^2]}\} \quad (20)$$

If we are to have conservation of momentum in any collisions of the kind considered, the function f must satisfy the equation (20) for all values of the independent variables u and v. The solution of this functional equation is obtained by letting u_1 in (20) tend to zero. In this way we get

$$f(v) = \frac{f(0)}{\sqrt{(1-v^2/c^2)}} \quad (21)$$

It is easily seen that the function $f(u)$ given by (21) satisfies (20) for all values of u_1 and v.

From (2) and (21) we thus obtain for the relativistic mass of a particle with the velocity $\mathbf{u}$

$$m = \frac{m_0}{\sqrt{(1-u^2/c^2)}}, \quad (22)$$

where we have put $$f(0) = m_0.$$

The constant m_0, the so-called proper mass or rest mass of the particle, is identical with the mass assigned to the particle in Newtonian mechanics, and the assumption preceding (10) obviously implies that the rest mass is unchanged in the collision considered. For the momentum of the particle we now get, according to (1) and (22),

$$\mathbf{p} = \frac{m_0\,\mathbf{u}}{\sqrt{(1-u^2/c^2)}}. \quad (23)$$

27. Force, work, kinetic energy

When the velocity of the particle and therefore its momentum are constant in time, this is taken as an indication that the particle is free. If, however, the momentum of the particle changes, the particle is said to be acted upon by a force **F** which is equal to the change of momentum per unit time:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}. \tag{24}$$

Equation (24), which for small particle velocities is identical with Newton's second law of mechanics, should here be regarded as the definition of a force in the relativistic region. It may be considered as an equation of motion only when it has been stated how **F** depends on the physical state of the system which is the cause of the change of momentum of the particle.

As in Newtonian mechanics, the work A done by the force per unit time is defined by

$$A = (\mathbf{F}\,.\,\mathbf{u}), \tag{25}$$

where $\mathbf{u}$ is the velocity of the particle. Further, the kinetic energy T of the particle is defined by the equation

$$\frac{dT}{dt} = A = (\mathbf{F}\,.\,\mathbf{u}), \tag{26}$$

expressing that the change of the kinetic energy per unit time is equal to the work A.

Using (23) and (24) the right-hand side of equation (26) may be written

$$A = \left(\mathbf{u}\,.\,\frac{d}{dt}\frac{m_0\,\mathbf{u}}{(1-u^2/c^2)^{\frac{1}{2}}}\right) = \frac{m_0}{(1-u^2/c^2)^{\frac{1}{2}}}\left(\mathbf{u}\,.\,\frac{d\mathbf{u}}{dt}\right) + \frac{m_0\,u^2}{c^2(1-u^2/c^2)^{\frac{3}{2}}}u\frac{du}{dt}$$

or, since

$$\left(\mathbf{u}\,.\,\frac{d\mathbf{u}}{dt}\right) = u\frac{du}{dt},$$

$$A = \frac{m_0\,u}{(1-u^2/c^2)^{\frac{3}{2}}}\frac{du}{dt} = \frac{d}{dt}\left(\frac{m_0\,c^2}{(1-u^2/c^2)^{\frac{1}{2}}}\right). \tag{27}$$

Introducing this expression into the right-hand side of (26), this equation can at once be integrated and we obtain for the kinetic energy of a particle with the velocity u

$$T = \frac{m_0\,c^2}{\sqrt{(1-u^2/c^2)}} + C, \tag{28}$$

where C is a constant of integration. Since the kinetic energy may be taken as zero for $u = 0$, we get

$$C = -m_0\,c^2,$$

and thus

$$T = \frac{m_0\,c^2}{\sqrt{(1-u^2/c^2)}} - m_0\,c^2. \tag{29}$$

If u is small compared with c, we can make an expansion in terms of $(u/c)^2$ and we then obtain, to a first approximation, the Newtonian expression for the kinetic energy

$$T = \tfrac{1}{2}m_0 u^2. \tag{30}$$

In the same way, all quantities of relativistic mechanics and the relations between these quantities are identical with the corresponding quantities and relations of Newtonian mechanics in the domain where the velocities are small compared with c. It is interesting to note that all deviations between relativistic and Newtonian mechanics are at least of second order in u/c, which explains why the early electron theory which was based on Newtonian mechanics was able to explain all effects of first order. When the velocity u approaches c, however, the deviations between relativistic mechanics and Newtonian mechanics are very large. For $u \to c$ both the mass (22) and the kinetic energy (29) become infinite, showing that in mechanics also c plays the role of a limiting velocity.

28. Transformation equations for momentum energy and force

Consider again the two systems of inertia S and S' corresponding to the special Lorentz transformation (II. 24). The momentum and the kinetic energy of a particle are then given in both systems by formulae of the form (23) and (29). From the transformation equations (II. 45) for particle velocities we then also obtain the transformation equations for momentum and energy of a particle.

Now it is convenient to introduce a quantity E defined by the equation

$$E = T+m_0 c^2 = \frac{m_0 c^2}{(1-u^2/c^2)^{\frac{1}{2}}} = mc^2. \tag{31}$$

The quantity E, differing from T by the constant amount $m_0 c^2$, is a measure of the kinetic energy of the particle and frequently we speak simply of the energy E of a free particle without, however, at the moment attributing any physical meaning to the constant $m_0 c^2$. A similar quantity $E' = T'+m_0 c^2$ is introduced in the system S'.

By means of (23), (31), and the relations inverse to (II. 45) and to (II. 48) we then get

$$p_x = \frac{m_0 u_x}{\sqrt{(1-u^2/c^2)}} = \frac{m_0(1+vu'_x/c^2)}{\sqrt{(1-v^2/c^2)}\sqrt{(1-u'^2/c^2)}}\left(\frac{u'_x+v}{1+vu'_x/c^2}\right)$$

$$= \frac{m_0(u'_x+v)}{\sqrt{(1-v^2/c^2)}\sqrt{(1-u'^2/c^2)}},$$

i.e.

$$p_x = \frac{p'_x+vE'/c^2}{\sqrt{(1-v^2/c^2)}}. \tag{32a}$$

In the same way we obtain

$$p_y = p'_y, \qquad p_z = p'_z, \qquad E = \frac{E'+vp'_x}{\sqrt{(1-v^2/c^2)}}. \tag{32 b}$$

A comparison of the equations (32) with the Lorentz transformations (II. 24′) shows that the four quantities

$$p_x, \quad p_y, \quad p_z, \quad E/c^2 \tag{33}$$

are transformed in the same way as the space-time coordinates x, y, z, t. Thus we have, in analogy to the equations (15), (13), and (13′) in Chapter II,

$$p^2-\frac{E^2}{c^2} = p'^2-\frac{E'^2}{c^2}. \tag{34}$$

From (23) and (31) we see that the invariant (34) has the constant value $-m_0^2c^2$. Therefore we have in any system of inertia

$$p^2-\frac{E^2}{c^2} = -m_0^2c^2,$$

i.e.

$$E = c(m_0^2c^2+p^2)^{\frac{1}{2}}. \tag{35}$$

Hence, for the velocity of the particle

$$u = \frac{pc^2}{E} = \frac{dE}{dp}. \tag{36}$$

Since the quantities (33) also transform like the coordinates (x, y, z, t) by rotations of the Cartesian axes, the transformation equations for momentum and energy may, in the case of the more general Lorentz transformation (II. 25′), be written in the vector form

$$\left.\begin{aligned} \mathbf{p} &= \mathbf{p}'+\frac{\mathbf{v}}{v^2}\frac{(\mathbf{v}\cdot\mathbf{p}')\{1-(1-v^2/c^2)^{\frac{1}{2}}\}+E'v^2/c^2}{\sqrt{(1-v^2/c^2)}} \\ E &= \frac{E'+(\mathbf{v}\cdot\mathbf{p}')}{\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}. \tag{37}$$

It is immediately seen that (34), which is analogous to (II. 30), is a consequence of (37).

Let us now consider a system of n free particles. If the momentum and energy of the ith particle are denoted by $\mathbf{p}^{(i)}$ and $E^{(i)} = T^{(i)}+m_0^{(i)}c^2$, respectively, where $T^{(i)}$ is the kinetic energy and $m_0^{(i)}$ is the rest mass of the ith particle, the total momentum and energy of the system are defined by

$$\left.\begin{aligned} \mathbf{p} &= \sum_{i=1}^{n}\mathbf{p}^{(i)}, \qquad E = \sum_i E^{(i)} = T+m_0c^2 \\ T &= \sum_i T^{(i)}, \qquad m_0 = \sum_i m_0^{(i)} \end{aligned}\right\}. \tag{38}$$

Since the transformation equations (32) or (37) are valid for each particle separately and as they are linear transformations, it is obvious that the same equations are valid also for the total momentum and energy of the system. Thus we can take over the equations (32) and (37) for a system of free particles, where $\mathbf{p}$, E, and T denote the total momentum and energy, while m_0, according to (38), is the sum of the proper masses of the particles. From the transformation equations (32) and (37) it is immediately seen that, if the theorem of conservation of momentum in a collision between the particles is valid in every system of inertia, the total energy E must also be conserved in any system of inertia.

Returning to a system consisting of a single particle, the equations (24) and (26) may be written in the form

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}, \tag{39a}$$

$$\frac{dE}{dt} = (\mathbf{F}\,.\,\mathbf{u}). \tag{39b}$$

Since the equations (39) are to hold in any system of inertia we can deduce the transformation equations for the force $\mathbf{F}$ from the known transformation properties of the quantities on the left-hand side of (39) From the equations (II. 25′) and (II. 26) we get

$$\frac{dt}{dt'} = \frac{1+(\mathbf{v}\,.\,\mathbf{u}')/c^2}{\sqrt{(1-v^2/c^2)}}.$$

Thus, by means of (37),

$$\mathbf{F} = \frac{d\mathbf{p}}{dt'}\frac{dt'}{dt}$$

$$= \frac{(1-v^2/c^2)^{\frac{1}{2}}\mathbf{F}' + \mathbf{v}\{(\mathbf{v}\,.\,\mathbf{F}')[1-\sqrt{(1-v^2/c^2)}] + (\mathbf{F}'\,.\,\mathbf{u}')v^2/c^2\}/v^2}{1+(\mathbf{v}\,.\,\mathbf{u}')/c^2}, \tag{40}$$

where $\mathbf{F}'$ is the force in the system S'. In relativistic mechanics the concept of force has no longer any absolute meaning as it has in Newtonian mechanics.

If we introduce the proper time τ of the particle instead of t and the Minkowski force

$$\mathbf{F}_M = \frac{\mathbf{F}}{\sqrt{(1-u^2/c^2)}}, \tag{41}$$

the equations (39) can, by means of (II. 38), be written

$$\frac{d\mathbf{p}}{d\tau} = \mathbf{F}_M, \qquad \frac{dE}{d\tau} = \mathbf{F}_M\,.\,\mathbf{u}. \tag{42}$$

Since τ is an invariant, the quantities $\{\dot{\mathbf{p}}, \dot{E}/c^2\}$ on the left-hand sides of (42) transform like the space-time coordinates $\{\mathbf{x}, t\}$: the same must be true, therefore, for the quantities $\{\mathbf{F}_M, (\mathbf{F}_M \cdot \mathbf{u})/c^2\}$. This can also be shown directly by an elementary calculation from (40), (II. 55), and (II. 56).

29. Hyperbolic motion. Motion of an electrically charged particle in a constant magnetic field

As mentioned before, the equation (39a) can be regarded as equation of motion only when it is known how the force $\mathbf{F}$ depends on the variables of the physical system causing the change of momentum of the particle. If the velocity of the particle is small relative to c, the relativistic equation must, however, be identical with Newton's second law and, in the inertial system S^0 relative to which the particle has velocity zero at the moment considered, we may assume that the force $\mathbf{F}^0$ is identical with the Newtonian force. By means of the transformation equations (40) we can then calculate the force $\mathbf{F}$ in an arbitrary inertial system S. Let the velocity of the particle relative to S be $\mathbf{u}$, if S' in (40) is the rest system S^0, we have $\mathbf{v} = \mathbf{u}$ and $\mathbf{u}' = 0$ and we then get for the force $\mathbf{F}$ in S

$$\mathbf{F} = \mathbf{F}^0(1-u^2/c^2)^{\frac{1}{2}}+\mathbf{u}\,\frac{(\mathbf{u}\,.\,\mathbf{F}^0)}{u^2}\{1-(1-u^2/c^2)^{\frac{1}{2}}\}, \tag{43}$$

where $\mathbf{F}^0$ is the Newtonian force. Decomposing the forces into two components respectively, parallel and perpendicular to $\mathbf{u}$, thus putting

$$\mathbf{F} = \mathbf{F}_{\parallel}+\mathbf{F}_{\perp}, \qquad \mathbf{F}^0 = \mathbf{F}^0_{\parallel}+\mathbf{F}^0_{\perp},$$

(43) can obviously be replaced by the simple equations

$$\mathbf{F}_{\parallel} = \mathbf{F}^0_{\parallel}, \qquad \mathbf{F}_{\perp} = \mathbf{F}^0_{\perp}(1-u^2/c^2)^{\frac{1}{2}}. \tag{43'}$$

If we know the Newtonian force $\mathbf{F}^0$ we can thus, by means of (43) or (43′), calculate $\mathbf{F}$ in any system S. In this way we can show, for example (Chap. V, § 58), that the force acting on an electrically charged particle travelling with the velocity $\mathbf{u}$ through an electric field $\mathbf{E}$ and a magnetic field $\mathbf{H}$ is given by the Lorentz formula

$$\mathbf{F} = e\left[\mathbf{E}+\frac{1}{c}(\mathbf{u}\times\mathbf{H})\right], \tag{44}$$

where e is the electric charge of the particle and $\mathbf{u}\times\mathbf{H}$ denotes the vector product of the vectors $\mathbf{u}$ and $\mathbf{H}$.

Now the left-hand side of the equation of motion (39a) can, by means of (1), (31), and (39b), be written

$$\frac{d\mathbf{p}}{dt} = \frac{d(m\mathbf{u})}{dt} = m\frac{d\mathbf{u}}{dt}+\frac{dm}{dt}\mathbf{u} = m\frac{d\mathbf{u}}{dt}+\frac{1}{c^2}\frac{dE}{dt}\mathbf{u} = m\frac{d\mathbf{u}}{dt}+\left(\frac{\mathbf{F}\,.\,\mathbf{u}}{c^2}\right)\mathbf{u},$$

where m is the relativistic mass given by (22). Introducing this into (39 a) we get the equations of motion in the form

$$m\frac{d\mathbf{u}}{dt} = \mathbf{F} - \frac{(\mathbf{F}.\mathbf{u})}{c^2}\mathbf{u}. \tag{45}$$

From this we see that the acceleration of the particle in general has a direction different from the direction of the force and, therefore, the motion of the particle is more complicated in relativistic mechanics than it is in Newtonian mechanics. Only if the force is constantly parallel or constantly perpendicular to **u**, will the motion of the particle again be relatively simple. We shall here treat an example of each of these cases.

Let us first consider a particle which is acted upon by a constant force $F = m_0 g$ and which has an initial velocity in the direction of the force. According to (45) the particle will then continue to move in the direction of the force. Therefore the path of the particle will be a straight line and we can choose this line as x-axis. (39 a) then reduces to

$$\frac{d}{dt}\left(\frac{u}{\sqrt{(1-u^2/c^2)}}\right) = \frac{F}{m_0} = g, \qquad u = \frac{dx}{dt}.$$

If we assume that the velocity is zero at the time $t = 0$, we get by a first integration

$$\frac{u}{\sqrt{(1-u^2/c^2)}} = gt$$

or

$$u = \frac{dx}{dt} = \frac{gt}{\sqrt{[1+(gt/c)^2]}}. \tag{46}$$

If we further assume that $x = 0$ when $t = 0$, we obtain by a second integration

$$x = \frac{c^2}{g}\left\{\left[1+\left(\frac{gt}{c}\right)^2\right]^{\frac{1}{2}} - 1\right\} \tag{47}$$

or

$$\left(x+\frac{c^2}{g}\right)^2 - c^2t^2 = \frac{c^4}{g^2}. \tag{48}$$

If we plot this motion in an xt-diagram the equation (48) will represent a hyperbola and, consequently, this motion is called a hyperbolic motion.†

As long as $(gt)^2 \ll c^2$, we can neglect powers of (gt/c) higher than the second, and we obtain from (47) the usual equation

$$x = \tfrac{1}{2}gt^2$$

representing the motion of a particle with constant acceleration.

† M. Born, *Ann. d. Phys.* **30**, 1 (1909); A. Sommerfeld, ibid. **33**, 670 (1910).

For large t, i.e. for large velocities, on the other hand, the increase in x with increasing t is slower than that according to Newtonian mechanics. For $t \to \infty$ the velocity u given by (46) approaches the finite value c independently of the value of g; thus, even if a particle is attacked by a very large constant force, it will never attain a velocity exceeding the velocity of light. This is in agreement with the considerations in § 21, p. 53. An electrically charged particle moving in a constant electric field with a velocity parallel to the direction of the field represents a case of the type just considered.

Let us now consider the motion of a charged particle in a constant *magnetic field* **H**. Decomposing the velocity **u** of the particle into components parallel and perpendicular to **H** respectively, thus putting $\mathbf{u} = \mathbf{u}_{\parallel} + \mathbf{u}_{\perp}$, the force acting on the particle may, according to (44), be written

$$\mathbf{F} = \frac{e}{c}(\mathbf{u}\times\mathbf{H}) = \frac{e}{c}(\mathbf{u}_{\perp}\times\mathbf{H}), \tag{49}$$

i.e. the force is perpendicular both to $\mathbf{u}$, $\mathbf{u}_{\perp}$, and $\mathbf{H}$. Therefore,

$$(\mathbf{F}\,.\,\mathbf{u}) = 0,$$

and (45) assumes the same form as in Newtonian mechanics, i.e.

$$m\frac{d\mathbf{u}}{dt} = \mathbf{F} = \frac{e}{c}(\mathbf{u}_{\perp}\times\mathbf{H}). \tag{50}$$

Here m denotes the relativistic mass which, however, in our case is constant, since E and thus also m and u, according to (39 b), are constant in time.

From (50) it follows that the component of the acceleration in the direction of **H** is zero, i.e. the component of the velocity in the direction of the field, viz. $u_{\parallel}$, is constant in time and, since also $u = |\mathbf{u}|$ is constant, the same must be the case for the magnitude $|\mathbf{u}_{\perp}|$ of the vector $\mathbf{u}_{\perp}$. Therefore the path of the particle must be a helix having its axis in the direction of the field The projection of the path on a plane perpendicular to **H** will be a circle with radius ρ determined by the condition that the centripetal force $mu_{\perp}^2/\rho$ in the circular motion must be equal to the force (49). Therefore, we have

$$\frac{mu_{\perp}^2}{\rho} = \frac{e}{c}u_{\perp}H$$

or

$$p_{\perp} = mu_{\perp} = \frac{e}{c}H\rho. \tag{51}$$

If the velocity of the particle is perpendicular to the direction of the field we get simply

$$p = \frac{e}{c}H\rho. \tag{52}$$

This equation enables us to determine the momentum of a charged particle by measuring **H** and ρ. This method has been especially useful in cosmic ray analysis and in β-ray and mass spectroscopy.

30. Equivalence of energy and mass

Let us again consider the system Σ_1 of n free particles. If $(\mathbf{p}, E)$ and $(\mathbf{p}', E')$ denote the total momentum and energy of the system in the two systems of inertia S and S', respectively, the connexion between the primed and unprimed variables is then given by (32) or (37). The invariant (34) will for such a system always have a negative value. For $n = 1$ this invariant is, according to (35), equal to $-m_0^2c^2$ and, keeping in mind that (35) holds for each particle, it is immediately seen that p^2-E^2/c^2 must be less than $-\sum_i (m_0^{(i)}c)^2$ for $n > 1$. This means that we can always choose an inertial system S' such that the total momentum $\mathbf{p}'$ of Σ_1 is zero in S'. Viz., putting $\mathbf{p}' = 0$ in (37), we obtain for the relative velocity $\mathbf{v}$ of S' and S

$$\mathbf{v} = \frac{c^2}{E}\mathbf{p} \tag{53}$$

and, since $p^2-E^2/c^2 < 0$, we have $E > cp$ so that the relative velocity v determined by (53) is less than c, which must be required.

Such a system of inertia S^0 in which the total momentum $\mathbf{p}^0 = 0$ is called the rest system of Σ_1 or the centre of gravity system since the system Σ_1 as a whole has the same mechanical properties as a particle at rest relative to S^0. Now let $\mathbf{u}$ be the velocity of S^0 relative to an arbitrary system of inertia S, then $\mathbf{u}$ also denotes the velocity with which the system Σ_1 as a whole moves relative to S.

If the system S' in (37) is identical with the rest system S^0, we get the equations

$$\mathbf{p} = \frac{E^0\mathbf{u}}{c^2\sqrt{(1-u^2/c^2)}}, \qquad E = \frac{E^0}{\sqrt{(1-u^2/c^2)}}, \tag{54}$$

which express the total momentum and energy of Σ_1 as a function of the velocity $\mathbf{u}$ of the system. It now seems natural to define the total mass of the system as the ratio between momentum and velocity in the same way as for a single particle (cf. equation (1)). From (54) we then see that we must assign to Σ_1 a mass M given by

$$M = \frac{E^0/c^2}{\sqrt{(1-u^2/c^2)}} = \frac{E}{c^2} = \frac{M_0}{\sqrt{(1-u^2/c^2)}}. \tag{55}$$

This corresponds to a rest mass $M_0 = E^0/c^2$ which, by means of (38), can be written in the form

$$M_0 = E^0/c^2 = m_0 + T^0/c^2, \tag{56}$$

where m_0 is the sum of the rest masses of all particles, and T^0 denotes the total kinetic energy of the particles in the rest system.

With the expression (56) for the rest mass of Σ_1 the equations (54) for the momentum and energy of Σ_1 become completely analogous to the equations (23) and (31) holding for a single particle. From (56) we see that the rest mass of Σ_1 is larger than the sum m_0 of the rest masses of all particles by the amount T^0/c^2. Thus the inner kinetic energy of the system contributes to the inertial mass of the system by an amount equal to T^0/c^2.

The invariant in (34) can, by means of (54) and (56), be written

$$p^2 - E^2/c^2 = -M_0^2 c^2 \tag{57}$$

in analogy to the equation (35) valid for a single particle. Equation (57) may also be used as a definition of the rest mass of Σ_1.

The important conclusion that the inner kinetic energy of a system of free particles corresponds to an inertial mass can now be extended to any kind of energy.

Starting from the assumed general validity of the conservation theorems,

$$\Delta m = \frac{\Delta E}{c^2}. \tag{58}$$

To prove this important theorem consider a 'collision' of the system Σ_1 of free particles with another arbitrary physical system Σ_2 by which a certain amount of energy and momentum is transferred from Σ_1 to Σ_2. Before and after this collision process the particles in Σ_1 are free; therefore the total momentum and energy of Σ_1 before and after the process are transformed according to equations of the form (32) and (37). Subtracting now the transformation equations for the energy and momentum after the 'collision' from the corresponding transformation equations before the collision, we get if $(\Delta \mathbf{p}, \Delta E)$ and $(\Delta \mathbf{p}', \Delta E')$ denote the difference between the momentum and energy of Σ_1 before and after the process measured in S and S', respectively,

$$\left.\begin{aligned} \Delta \mathbf{p} &= \Delta \mathbf{p}' + \frac{\mathbf{v}}{v^2}\,\frac{(\mathbf{v}.\Delta \mathbf{p}')\{1-(1-v^2/c^2)^{\frac{1}{2}}\}+v^2\,\Delta E'/c^2}{\sqrt{(1-v^2/c^2)}} \\ \Delta E &= \frac{\Delta E' + (\mathbf{v}.\Delta \mathbf{p}')}{\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}. \tag{59}$$

If momentum and energy are to be conserved during the process the momentum and energy of the system Σ_2 must be increased by the amounts $\Delta\mathbf{p}$ and ΔE in S and by $\Delta\mathbf{p}'$ and $\Delta E'$ in S'.

In analogy to (34) it now follows from (59) that the quantity

$$|\Delta\mathbf{p}|^2-\frac{(\Delta E)^2}{c^2}=|\Delta\mathbf{p}'|^2-\frac{(\Delta E')^2}{c^2} \tag{60}$$

is an invariant. When this invariant has a negative value, it is again possible to find a system of inertia $S'=S^0$ in which $\Delta\mathbf{p}'=\Delta\mathbf{p}^0=0$, and if $\mathbf{u}$ is the velocity of S^0 relative to S, we get from (59)

$$\Delta\mathbf{p}=\frac{\mathbf{u}\,\Delta E^0/c^2}{\sqrt{(1-u^2/c^2)}},\qquad \Delta E=\frac{\Delta E^0}{\sqrt{(1-u^2/c^2)}}. \tag{61}$$

A comparison of these equations with (22) and (23) shows that we must assign to the energy ΔE^0 transferred to Σ_2 a rest mass

$$\Delta m_0=\frac{\Delta E^0}{c^2} \tag{62}$$

and a mass Δm relative to S, which is

$$\Delta m=\frac{\Delta m_0}{(1-u^2/c^2)^{\frac{1}{2}}}=\frac{\Delta E^0/c^2}{(1-u^2/c^2)^{\frac{1}{2}}}=\frac{\Delta E}{c^2}; \tag{63}$$

for the change of momentum of Σ_2 during the process is the same as if a material particle with the rest mass Δm_0 and the velocity $\mathbf{u}$ relative to S had been added to the system Σ_2. The invariant (60) can now be written

$$|\Delta\mathbf{p}|^2-\frac{(\Delta E)^2}{c^2}=-(\Delta m_0)^2c^2 \tag{64}$$

in analogy to (35). This equation gives a simple expression for the rest mass of the transferred energy.

Since Σ_2 was a general physical system so that the transferred energy may have any form, we see that formula (63) must hold for any type of energy. The system Σ_2 may, for instance, be an electromagnetic field such that the energy transferred from Σ_1 has taken the form of electromagnetic radiation. Thus the transformation equations for the energy and momentum of electromagnetic radiation must be given by (59). Further, if Σ_2 is a body which transforms the absorbed energy ΔE into heat, we see also that the heat energy of a body contributes to its mass, so that the mass of a body is increased by heating. If, finally, Σ_2 were a system converting the transferred energy into potential energy, it follows that inertial mass has to be assigned also to the potential energy of a system.

From the preceding argument, in particular from equation (64), it follows that the notion of mass of a certain amount of energy ΔE has a well-defined meaning only if we also know the momentum $\Delta\mathbf{p}$ connected with the energy, and for this mass to be real the left-hand side of (64) must be negative. Only for

$$|\Delta E| > c|\Delta\mathbf{p}| \tag{65}$$

may we speak at all of a rest system and, consequently, of a certain velocity and of a real mass of the energy.

In the discussion above ΔE may be negative, so that we actually have to deal with a transfer of energy to the system Σ_1. Thus, considering a process in which the total momentum $\mathbf{p}_2$ and energy E_2 of Σ_2 are transferred to Σ_1, we see that the transformation equations (59) must hold for the total momentum and energy of an arbitrary system. If, furthermore, the relation

$$p_2^2 - \frac{E_2^2}{c^2} \leqslant 0 \tag{66}$$

holds, the system has a real rest mass defined by (64), i.e. by

$$-M_0^2 c^2 = p_2^2 - \frac{E_2^2}{c^2}. \tag{67}$$

In the case where $p_2^2 - E_2^2/c^2$ is equal to zero the rest mass of the system will also be zero.

If Σ_2 is a system of fields where the density of field energy is a homogeneous positive definite function of the field variables, (66) must always be fulfilled. If E_2 were less than cp_2, we could find a system of inertia S' in which $E_2' = 0$ and $\mathbf{p}_2' \neq 0$. For this purpose it would be sufficient to give $\mathbf{v}$ in the transformation equations (59) for $\mathbf{p}_2$ and E_2 the value

$$\mathbf{v} = E_2\,\mathbf{p}_2/p_2^2 \quad (|\mathbf{v}| < c). \tag{68}$$

According to the relations inverse to (59) we should then have

$$E_2' = \frac{E_2 - (\mathbf{v}\cdot\mathbf{p}_2)}{\sqrt{(1-v^2/c^2)}} = 0, \qquad p_2' = (p_2^2 - E_2^2/c^2)^{\frac{1}{2}} > 0.$$

According to the assumptions made above regarding the dependence of the energy density on the field variables, E_2' can, however, be zero only if the field itself vanishes and in this case the momentum of the field $\mathbf{p}_2'$ must also be zero. It is therefore natural to assume that the relation (66) holds for any macroscopic physical system. As a simple example of a system where the sign of equality holds in (66) we have the case of a train of electromagnetic plane waves. According to (67), such a system must therefore have the rest mass zero and a velocity $u = c^2 p_2/E_2 = c$ in any system of inertia.

Now it may also be shown that, conversely, any material particle with the mass m must represent an energy $E = mc^2$ and in particular in the rest system of the particle an energy $E_0 = m_0 c^2$. This statement obviously has a real meaning only when a process exists in which the energy represented by the mass of the particle can be transformed into another form of energy such as the kinetic energy of other particles. We cannot know in advance whether such an 'annihilation process' actually does exist in nature, but we can show that if it exists under certain conditions and if we require that the principle of relativity as well as the theorems of conservation of momentum and energy shall hold for this process, the amount of energy liberated by annihilation of the mass m_0 will be equal to $E_0 = m_0 c^2$.

To prove this statement, we assume that the momentum and energy liberated by the annihilation process are transferred to the previously considered system Σ_1 of free particles. Consider again two systems of inertia S and S^0, where S^0 is the rest system of the particle. If $(\Delta\mathbf{p}, \Delta E)$ and $(\Delta\mathbf{p}^0, \Delta E^0)$ denote the transferred momentum and energy measured in S and S^0, respectively, we have, according to (59)

$$\Delta\mathbf{p} = \Delta\mathbf{p}^0 + \frac{\mathbf{u}}{u^2}\,\frac{(\mathbf{u}\,.\,\Delta\mathbf{p}^0)\{1-\sqrt{(1-u^2/c^2)}\}+u^2\,\Delta E^0/c^2}{\sqrt{(1-u^2/c^2)}}, \tag{69}$$

where $\mathbf{u}$ is the velocity of the particle relative to S.

Since the momentum of the particle is zero in the rest system and since, moreover, momentum and energy of the particle are zero in any system of inertia after its annihilation, we obviously have

$$\Delta\mathbf{p}^0 = 0, \qquad \Delta E^0 = E_0, \tag{70}$$

where E_0 denotes the unknown energy content of the particle at rest before the annihilation. Furthermore, $\Delta\mathbf{p}$ is equal to the momentum $\mathbf{p}$ of the particle relative to S before the annihilation:

$$\Delta\mathbf{p} = \mathbf{p} = \frac{m_0\,\mathbf{u}}{\sqrt{(1-u^2/c^2)}}. \tag{71}$$

Substituting from (70) and (71) in (69) we obtain

$$\frac{m_0\,\mathbf{u}}{\sqrt{(1-u^2/c^2)}} = \frac{E_0\,\mathbf{u}/c^2}{\sqrt{(1-u^2/c^2)}},$$

or

$$E_0 = m_0 c^2. \tag{72}$$

The total energy content of a free particle with the velocity u must then be

$$E = E_0 + T = \frac{m_0 c^2}{\sqrt{(1-u^2/c^2)}} = mc^2, \tag{73}$$

where we have made use of (29) and (22). The quantity E introduced in (31) has thus a deeper physical meaning. The energy (72) is called the proper energy or rest energy of the particle.

We have thus obtained a general proof of Einstein's famous formula†

$$E = mc^2, \tag{74}$$

which states that any energy E has an inertia corresponding to a mass E/c^2, and that any mass m represents an energy mc^2. This theorem of the equivalence of energy and mass is one of the most important results of the special theory of relativity. It should be noted that the mass involved in this theorem is the inertial mass. One of the basic assumptions of the general theory of relativity is, however, (cf. Chap. VIII, § 83), the equality of inertial and gravitational mass, so that we shall have to ascribe to an energy of amount E also a gravitational mass m given by (74).

31. Inelastic collisions. Mass of a closed system of particles

Before proceeding to a discussion of the experimental verification of relativistic mechanics we shall, in the present section, consider a few simple examples illustrating the general theorem of the equivalence of mass and energy

Let us first consider a completely inelastic head-on collision between two clay balls of the same rest mass m_0 which in the inertial system $S' = S^0$ travel along the same line with velocities of equal magnitude but opposite in direction. The total momentum and energy in the centre of gravity system S^0 before the collision are then

$$\mathbf{p}^0 = 0, \qquad E^0 = 2m_0 c^2 + T^0, \tag{75}$$

where T^0 is the total kinetic energy. In a system of inertia S relative to which S^0 has the velocity $\mathbf{u}$, we thus have, according to (37) and (75), for the total momentum and energy before the collision

$$\mathbf{p} = \frac{(2m_0 + T^0/c^2)\mathbf{u}}{\sqrt{(1-u^2/c^2)}}, \tag{76a}$$

$$E = \frac{2m_0 c^2 + T^0}{\sqrt{(1-u^2/c^2)}}. \tag{76b}$$

By the inelastic collision, the two clay balls will unite and form one large ball; on account of the theorem of conservation of momentum applied in S^0, this ball will have zero momentum and thus also zero velocity in S^0.

† A Einstein, *Ann d Phys.* **18**, 639 (1905), **23**, 317 (1907), H. A. Lorentz, *Das Relativitatsprinzip* 3 Haarlemer Vortrage, Leipzig 1914, id, *Amst. Versl.* **20**, 87 (1911).

The original kinetic energy T^0 of the particles in S^0 is converted into heat and, on account of the conservation of energy, the amount of heat developed is

$$Q^0 = T^0. \tag{77}$$

The clay ball after the collision now has the velocity $\mathbf{u}$ relative to S, i.e. a momentum

$$\mathbf{p} = \frac{M_0 \mathbf{u}}{\sqrt{(1-u^2/c^2)}}, \tag{78}$$

where M_0 is the rest mass of the ball after the collision. As the conservation of momentum theorem must hold also in S, the expressions (78) and (76a) must be equal. Thus, applying (77), we get

$$M_0 = 2m_0 + \frac{Q^0}{c^2}, \tag{79}$$

i e. the rest mass of the ball M_0 after the collision is equal to the sum of the rest masses of the original balls increased by the mass of the heat energy.

For the energy of the ball after the collision in S, we obtain, according to (79) and (77),

$$E = \frac{M_0 c^2}{\sqrt{(1-u^2/c^2)}} = \frac{2m_0 c^2 + T^0}{\sqrt{(1-u^2/c^2)}}, \tag{80}$$

which is equal to the total energy (76b) of the balls before the collision, in accordance with the theorem of conservation of energy in S. For the difference between the total kinetic energy before and after the collision we obtain, however,

$$T_{\text{before}} - T_{\text{after}} = \left(\frac{2m_0 c^2 + T_0}{\sqrt{(1-u^2/c^2)}} - 2m_0 c^2\right) - M_0 c^2\left(\frac{1}{\sqrt{(1-u^2/c^2)}} - 1\right) = Q^0 \tag{81}$$

by means of (76b), (79), and (77). This difference is thus an invariant independent of the system of inertia in which the kinetic energies are calculated.

To give an illustration of the inertia of potential energy we now consider a system Σ consisting of a certain number of particles which are held together by attractive forces. Let us assume that there exist inertial systems S' in which all particle velocities are small compared with the velocity of light such that in S' we can use the non-relativistic Newtonian mechanics as a good approximation. Neglecting the typically atomic phenomena connected with the existence of Planck's quantum of action, we may consider an atomic nucleus to form such a mechanical system, the atomic nuclei being built up of elementary particles, the nucleons, which are so heavy that their velocities can generally be regarded as small compared with c. This assumption means that the proper times of the separate particles in Σ are practically identical and equal to the time t' in S' and, furthermore, that the forces between the particles can be considered to a first approximation to be functions of the positions of the particles. If, moreover, the forces are conservative forces, they can in S' be expressed as gradients of a potential function V' which is a function of the position coordinates of the particles. According to Newtonian mechanics the particles in the system will now move in such a way that the sum of the kinetic energies and the potential energy is constant in time, i.e.

$$T'+V' = H' = \text{constant.} \tag{82}$$

However, the total kinetic energy is, of course, not constant in general. We now adjust the arbitrary constant in the potential energy V' so that $V' = 0$ when the particles are so far from each other that the forces are zero. Thus V' is obviously negative in any state where the particles are bound to each other.

Among the possible systems of inertia S' we now choose that system S^0 in which the centre of gravity of Σ is at rest. The sum $\mathbf{p}^0$ of the momenta of the particles is then obviously constant and equal to zero in S^0. In a system of inertia S relative to which S^0 is moving with the velocity u in the direction of the x-axis, the sum $\mathbf{p}$ of the momenta of the particles is, according to (32), given by

$$p_x = \frac{(m_0+T^0/c^2)u}{\sqrt{(1-u^2/c^2)}}, \qquad p_y = p_z = 0, \tag{83}$$

where again m_0 is the sum of the rest masses of all the particles. Since the proper times of the particles of the system are practically identical, $\mathbf{p}$ can be regarded as a function of a well-defined time variable. In contrast to the case of a system of free particles, $\mathbf{p}$ is *not* constant in time since the kinetic energy T^0 occurring in the expression (83) for $\mathbf{p}$ is time-dependent.

However, $\mathbf{p}$ is not the total momentum of Σ in S, since we have to take into account the fact that the potential energy V^0 has a rest mass V^0/c^2. The total momentum $\mathbf{P}$ in S is therefore

$$P_x = p_x + \frac{V^0u/c^2}{\sqrt{(1-u^2/c^2)}} = \frac{(m_0+H^0/c^2)u}{\sqrt{(1-u^2/c^2)}}, \qquad P_y = P_z = 0, \tag{84}$$

where $H^0 = T^0+V^0$ is the total energy in the rest system. In contrast to $\mathbf{p}$, the total momentum vector $\mathbf{P}$ is obviously constant in time, as is required for a closed system with no external forces. If the particles are so far away from each other that they can be regarded as free, we must have $\mathbf{P} = \mathbf{p}$, a condition which determines uniquely the constant in the potential energy.

From (84) we see that an atomic nucleus which moves as a whole with the velocity $\mathbf{u}$ relative to a system of inertia S has a total momentum

$$\mathbf{P} = \frac{M_0 \mathbf{u}}{\sqrt{(1-u^2/c^2)}} \tag{85}$$

with a rest mass

$$M_0 = m_0 + H^0/c^2. \tag{86}$$

For stable nuclei H^0 is negative and $\Delta E = -H^0$ is the *binding energy* of the nucleus, i.e. the amount of energy which must be transferred to the nucleus in order to disintegrate it completely into its constituent particles. For the quantity $\Delta m = m_0 - M_0$, the *mass defect* of the nucleus, we have, according to (86),

$$\Delta m = \frac{\Delta E}{c^2}. \tag{87}$$

This fundamental relation connecting the binding energy and the mass defect, which is a special case of Einstein's formula (74), has now been verified experimentally with great accuracy in many nuclear reactions (cf. § 32).†

32. Experimental verification of relativistic mechanics

In § 26 we saw that, if the theorem of conservation of momentum is to hold in any elastic collision between two particles, the variation of the mass with velocity must be given by (22). The next question is whether something like conservation of momentum and energy exists at all for large velocities. This question can only be settled by experiments. Direct experiments of this kind were performed by Champion,‡ who investigated the collisions between rapidly moving electrons (β-particles) and electrons at rest in a Wilson chamber. Let us consider more closely such a collision between an electron 1 with velocity $\mathbf{u}_1$ and momentum $\mathbf{p}_1 = m_0 \mathbf{u}_1(1-u_1^2/c^2)^{-\frac{1}{2}}$ and an electron 2 with velocity zero, i.e. with the momentum $\mathbf{p}_2 = 0$ relative to a system of coordinates S in which the cloud chamber is at rest. In the collision a certain momentum is transferred to particle 2. Let $\bar{\mathbf{p}}_1$ and $\bar{\mathbf{p}}_2$ be the momenta of the two particles after the collision, and θ and ϕ the angles between the direction of the incident electron and the vectors $\bar{\mathbf{p}}_1$ and $\bar{\mathbf{p}}_2$, respectively (cf. Fig. 11).

† The possibility of such an effect was first discussed by P. Langevin, *J. de Phys.* (5), **3**, 553 (1913).

‡ F. C. Champion, *Proc. Roy. Soc.* A, **136**, 630 (1932).

Application of the conservation laws of relativistic mechanics then leads to a simple relation between θ and ϕ, which we shall now deduce.

It is convenient here to make use of the fact that the theorems of conservation of momentum and energy are valid in any system of inertia (if they hold at all). For convenience we choose the Cartesian axes in

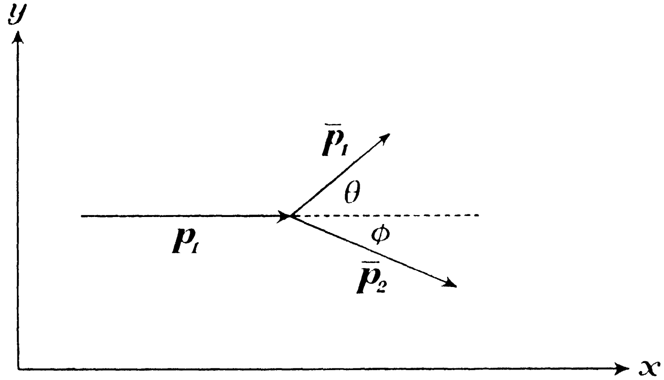

FIG. 11

the laboratory system S such that $\mathbf{p}_1$ is parallel to the x-axis and $\bar{\mathbf{p}}_1$ lies in the xy-plane. Then, if the momentum is conserved, $\bar{\mathbf{p}}_2$ also must lie in the xy-plane. We now introduce the centre of gravity system S' in which the total momentum $\mathbf{p}' = \mathbf{p}'_1+\mathbf{p}'_2 = 0$. S' is moving relative to S with a certain velocity v in the direction of the x-axis. From the reciprocal equation to (32 a)

$$p'_x = \frac{p_x - vE/c^2}{\sqrt{(1-v^2/c^2)}},$$

which holds for each particle separately as well as for the total momentum $\mathbf{p}$ and energy E of the system, we then get, since $\mathbf{p} = \mathbf{p}_1$ and $E = E_1+m_0c^2 = m_0c^2+m_0c^2(1-u_1^2/c^2)^{-\frac{1}{2}}$, the following expression for v:

$$v = \frac{c^2p_1}{E} = \frac{c^2m_0u_1}{\sqrt{(1-u_1^2/c^2)}m_0c^2\{1+(1-u_1^2/c^2)^{-\frac{1}{2}}\}} = \frac{u_1}{1+\sqrt{(1-u_1^2/c^2)}}. \tag{88}$$

Since $\mathbf{p}'_2 = -\mathbf{p}'_1$, the two particles have the same initial velocity u' relative to S'. Since, furthermore, particle 2 was at rest in S, the velocity u' relative to S' must be equal to the relative velocity v of S' and S, i.e.

$$u' = v, \qquad \tfrac{1}{2}E' = E'_1 = E'_2 = \frac{m_0c^2}{\sqrt{(1-v^2/c^2)}}. \tag{89}$$

Applying now the theorems of conservation of momentum and energy in the centre of gravity system, we obtain for the momenta and energies of the particles after the collision

$$\bar{\mathbf{p}}'_2 = -\bar{\mathbf{p}}'_1, \qquad \bar{E}'_1 = \bar{E}'_2 = \tfrac{1}{2}\bar{E}' = \frac{m_0 c^2}{\sqrt{(1-v^2/c^2)}}$$
$$\bar{p}'_2 = \bar{p}'_1 = \sqrt{(\bar{E}'^2_1/c^2 - m_0^2 c^2)} = \frac{m_0 v}{\sqrt{(1-v^2/c^2)}}. \tag{90}$$

Consequently the two particles will also have the same velocity v relative to S' after the collision and, according to (90), we have

$$\frac{v^2 \bar{E}'^2_1}{c^4} = \bar{p}'^2_1. \tag{91}$$

From Fig. 11 we see at once that

$$\tan\theta = \frac{\bar{p}_{1y}}{\bar{p}_{1x}}, \qquad \tan\phi = \frac{-\bar{p}_{2y}}{\bar{p}_{2x}}.$$

Using the transformation equation (32) for the momenta of the two particles after the collision, we obtain by means of (90) and (91)

$$\tan\theta\tan\phi = \frac{\bar{p}'_{1y}(-\bar{p}'_{2y})(1-v^2/c^2)}{(\bar{p}'_{1x}+v\bar{E}'_1/c^2)(\bar{p}'_{2x}+v\bar{E}'_2/c^2)} = \frac{\bar{p}'^2_{1y}(1-v^2/c^2)}{(v\bar{E}'_1/c^2+\bar{p}'_{1x})(v\bar{E}'_1/c^2-\bar{p}'_{1x})}$$
$$= \frac{\bar{p}'^2_{1y}(1-v^2/c^2)}{v^2\bar{E}'^2_1/c^4-\bar{p}'^2_{1x}} = 1-\frac{v^2}{c^2}. \tag{92}$$

Introducing the expression (88) for v into (92) we finally get

$$\tan\theta\tan\phi = \frac{2}{\gamma_1+1}, \tag{93}$$

where we have put $$\gamma_1 = (1-u_1^2/c^2)^{-\frac{1}{2}}. \tag{94}$$

Thus, the product of $\tan\theta$ and $\tan\phi$ is independent of the rest mass of the particles (as long as they are equal) and is a function of the velocity of the incident particle alone.

In the limit $c \to \infty$, we obtain the corresponding formula of Newtonian mechanics. In this limit,

$$\gamma_1 = 1 \quad \text{and} \quad \tan\theta\tan\phi = 1,$$

i.e. $$\tan(\theta+\phi) = \frac{\tan\theta+\tan\phi}{1-\tan\theta\tan\phi} = \infty$$

or $$\theta+\phi = \tfrac{1}{2}\pi. \tag{95}$$

According to Newtonian mechanics the directions of motion of the two particles after the collision should thus be perpendicular to each other. This well-known effect can be observed in collisions between billiard balls, for instance, and has also been verified in collisions between α-particles and helium nuclei in a Wilson chamber, the velocity of the

α-particles being so small compared with c that Newtonian mechanics may be applied in this case. When the velocity of the incident particle approaches the velocity of light, however, $\gamma_1 > 1$ and, thus,

$$\tan\theta\tan\phi < 1,$$

which shows that the angle between the directions of motion of the particles after the collision in this case will be smaller than $\frac{1}{2}\pi$.

This characteristic phenomenon is well suited to an experimental test of relativistic mechanics. Cloud-chamber pictures of the collision of β-particles with electrons at rest in the chamber allow a direct determination of the angles θ and ϕ. It turns out that in many cases $\theta+\phi$ is considerably smaller than 90°. If, in addition, the velocities, and consequently also γ_1, for the incident β-particles are known, we arrive at a direct test of the relativistic formula (93). Measurements of that kind were performed by Champion,† who found agreement with formula (93). This may be regarded as an experimental proof of the theorems of conservation of momentum and energy in collisions between electrons, and thus also of the relativistic formula (22) for the variation of mass with velocity. The accuracy of these measurements was quite sufficient to disprove the formula for the variation of mass with velocity put forward by Abraham‡ before the development of the theory of relativity.

The equations of motion (39*a*) and (44) can also be tested directly by measuring the deflexions of rapidly moving electrons in electric and magnetic fields. In the case of a constant magnetic field perpendicular to the direction of motion of the electron, the previously derived equation (52) should hold, i.e.

$$\frac{mu}{e} = \frac{H\rho}{c}. \tag{96}$$

If, on the other hand, the electron traverses a constant electric field perpendicular to the initial velocity $\mathbf{u}$ of the electron, the electron will, according to (45) and (44), have a constant acceleration $g = eE/m$ in the direction of the field as long as the deflexion from the original path is small. In the time $t = l/u$ which the electron requires to traverse a distance l in the field, the electron is thus deflected a distance

$$\Delta l = \tfrac{1}{2}gt^2 = \frac{1}{2}\frac{eE}{m}\frac{l^2}{u^2} \tag{97}$$

perpendicular to the initial direction of motion.

If, therefore, the electron traverses both a magnetic and an electric field of the above-mentioned kind, simultaneous measurements of H, ρ, E, l,

† See ref., p. 85.

‡ M. Abraham, *Ann. d. Phys.* **10**, 105 (1903).

and Δl allow us to determine the variation of m with u by means of (96) and (97). Such experiments were performed by a large number of investigators† and the results obtained in the later experiments are in close agreement with the relativistic formula (22). Simultaneously the experiments may be regarded as a test of Lorentz's expression (44) for the force acting on an electrically charged moving particle.

If we know the charge of the particle from other experiments, measurements of the deflexion of the particle in known electric and magnetic fields allow us to determine the absolute value of the mass of the particle. This is used in the so-called mass spectrograph, which permits a very precise determination of the mass of atomic nuclei.

While thus the relativistic equations of motion of fast charged particles were experimentally proved at an early time, it was not until recent years possible to test the equations (74) and (87) which express the equivalence of energy and mass. This is understandable if we bear in mind that the change in mass of a body due to its potential energy or due to heating according to these equations in general will be negligible compared with the total mass of the body. The situation is different, however, when Einstein's relation is applied to single atomic nuclei.

From the results obtained at the end of the preceding section we should expect the mass of an atomic nucleus in its ground state to be always smaller than the sum of the masses of the constituent nucleons. According to (87) this mass defect should be particularly large for the stablest nuclei with large binding energy. Measurements of nuclear masses by means of the mass spectrograph have confirmed this result and in some cases the mass defect amounts to several per cent. of the mass of the whole nucleus.

As soon as one succeeded in producing nuclear reactions in which nuclei with a given mass defect are transformed into nuclei with other mass defects it also became possible to test Einstein's formula (74) experimentally. Here we shall only discuss one single example of such a process. When lithium is bombarded by fast protons (Cockcroft and Walton, 1932‡) it happens that a proton ($_1^1$H) penetrates into a lithium nucleus ($_3^7$Li), thus forming a compound nucleus which is unstable and quickly splits into two fast α-particles ($_2^4$He). This process can be

† Among the numerous papers we quote W. Kaufmann, *Gott Nachr. math.-nat. Klasse*, p. 143 (1901); A. H. Bucherer, *Verh. d. Deutschen Phys. Ges.* **6**, 688 (1908), G. Neumann, *Ann. d. Phys.* **45**, 529 (1914); Ch. E. Guye and Ch. Lavanchy, *Arch. de Genève*, **41**, 286, 353, 441 (1916). See also the comprehensive account by W Gerlach in *Hdb. d. Phys* **22**, 61 (1926).

‡ J. D. Cockcroft and G. T. S. Walton, *Proc. Roy. Soc.* A, **137**, 229 (1932).

expressed by the equation

$$ {}^{7}_{3}\mathrm{Li}+{}^{1}_{1}\mathrm{H} = {}^{4}_{2}\mathrm{He}+{}^{4}_{2}\mathrm{He}. \tag{98}$$

The masses of the nuclei entering in this process are known very well from mass-spectrograph measurements. According to the latest measurements the mass of ${}^{7}_{3}\mathrm{Li}$ is 7·0166 if the mass of the oxygen atom is put equal to 16. Now, if we also content ourselves with four decimal places for the hydrogen nucleus and the helium nucleus, the corresponding figures for ${}^{1}_{1}\mathrm{H}$, ${}^{4}_{2}\mathrm{He}$ are 1·0076 and 4·0028, respectively.† Thus the loss in mass m during the process is

$$m = 7{\cdot}0166+1{\cdot}0076-2\times 4{\cdot}0028 = 0{\cdot}0186$$

mass units, or in grammes,

$$m = 0{\cdot}309\times 10^{-25}\ \mathrm{gm}. \tag{99}$$

According to Einstein's relation this mass represents an energy of

$$mc^2 = 27{\cdot}7\times 10^{-6}\ \mathrm{erg}, \tag{100}$$

which after the process must appear as kinetic energy of the α-particles By means of the range-energy relation the kinetic energy of the particles can now be determined from measurements of the ranges of the α-particles. Newer very precise measurements‡ give for the difference between the total kinetic energy of the α-particles and the kinetic energy of the incident proton the value of

$$E = 17{\cdot}28\pm 0{\cdot}03\ \mathrm{MeV} = (27{\cdot}6\pm 0{\cdot}05)\times 10^{-6}\ \mathrm{erg}. \tag{101}$$

The agreement between (100) and (101) is excellent. The deviation is less than the possible error in the accuracy of the measurements, the error in the determination of the mass of the lithium nucleus involving an error in mc^2 of $\pm 0{\cdot}2\times 10^{-6}$. Einstein's equation (74) can thus be regarded as being verified experimentally with high accuracy, the error being less than 1 per cent.

While the mass corresponding to the heat developed during a usual chemical process is immeasurably small, the amount of heat developed in a nuclear reaction pile is so large that the corresponding mass may be of the order of many grammes. In the fission of a uranium nucleus, where the nucleus is divided into two fragments, an energy ΔE is developed whose mass $\Delta m = \Delta E/c^2$ amounts to about a thousandth of the mass of the uranium nucleus. By complete fission of all the nuclei in 1 ton of uranium the mass of the heat developed would thus be of the order of magnitude of 1 kg.

† K. T. Bainbridge and K. B. Jordan, *Phys. Rev.* **51**, 384 (1937), see also H. Bethe and M. S. Livingston, *Rev. Mod. Phys* **9**, 370 (1937).

‡ N. M. Smith, Jr., *Phys. Rev* **56**, 548 (1939).

Modern atomic physics, on the other hand, could also verify the equation (72) according to which a particle of mass m_0 represents an energy $E_0 = m_0 c^2$. This statement can be tested only (and therefore has a real meaning only) if a process exists in which the particle is annihilated completely. After the discovery of the positive electrons, the positons, in 1932† it became clear that such annihilation processes do exist in which a positive and a negative electron (a positon and a negaton) are annihilated, in accordance with Dirac's theory of electrons.‡ Since both particles have the same mass m_0, the amount of energy liberated in such a process should, according to (72), be equal to $2m_0 c^2$. This energy is emitted in the form of electromagnetic radiation and by measurements of the energy of this radiation equation (72) could be verified. The reverse process, in which electromagnetic radiation (light quanta) is transformed into pairs of positive and negative electrons, is also possible provided the light quanta participating in each individual process have the necessary energy of $2m_0 c^2$.§

Summarizing, we can state that all consequences of relativistic mechanics have proved to be in complete agreement with the measurements. The objection which could be raised against this statement is that all these measurements were performed on atomic particles, electrons and nuclei, etc., which in view of the existence of the quantum of action cannot be treated by means of classical mechanics. Since, on the other hand, classical mechanics is only a special case of quantum mechanics, the experiments mentioned above simply show that such consequences as the theorem of the equivalence of energy and mass must be valid beyond the domain of classical mechanics. In the general proof of this theorem, given in § 30, the system Σ_2 was in fact a quite general physical system without any limiting assumptions regarding its constitution. Moreover, in all experiments on the deflexion of electrons in a macroscopic electromagnetic field we are concerned with a case where quantum mechanics goes over into classical mechanics so that these experiments can be regarded as a direct verification of the fundamental equations of classical relativistic point mechanics.

† C. D. Anderson, *Science*, **76**, 238 (1932), *Phys. Rev.* **43**, 491 (1933), P. M. S. Blackett and G. P. S. Occhialini, *Proc. Roy. Soc.* A, **139**, 699 (1933).

‡ P. A. M. Dirac, *The Principles of Quantum Mechanics*, 3rd ed., Oxford 1947, § 73.

§ C. D. Anderson and S. H. Neddermeyer, *Phys. Rev.* **43**, 1034 (1933). See also F. Rasetti, L. Meitner, and K. Philipp, *Naturw.* **21**, 286 (1933), J. Curie and F. Joliot, *C.R.* **196**, 1581 (1933).

*

IV

FOUR-DIMENSIONAL FORMULATION OF THE THEORY OF RELATIVITY: TENSOR CALCULUS

33. Four-dimensional representation of the Lorentz transformation

In a definite system of inertia S an arbitrary event is characterized by the four space-time coordinates (x, y, z, t). In another system of inertia S' the *same* event is characterized by four other numbers (x', y', z', t'). If we assume that the origins of the Cartesian coordinates in the two systems S and S' coincide at the time $t = t' = 0$, the connexion between these space-time coordinates is given by a homogeneous Lorentz transformation, i.e. by a homogeneous linear transformation leaving the quantity s^2 (II. 13) invariant, i.e.

$$s^2 = x^2+y^2+z^2-c^2t^2 = x'^2+y'^2+z'^2-c^2t'^2 = s'^2. \tag{1}$$

Putting

$$\begin{aligned} x_1 &= x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict \\ x_1' &= x', \quad x_2' = y', \quad x_3' = z', \quad x_4' = ict', \end{aligned} \tag{2}$$

where $i = \sqrt{(-1)}$ is the imaginary unit, the homogeneous Lorentz transformation can thus be characterized as a homogeneous linear transformation

$$x_i' = \sum_{k=1}^{4} \alpha_{ik} x_k \quad (i = 1, 2, 3, 4), \tag{3}$$

which satisfies the condition

$$s^2 = \sum_i x_i^2 = \sum_i x_i'^2. \tag{4}$$

Here the coefficients α_{ik} depend only on the angles between the spatial axes in S and S' and on the relative velocity of the two systems of inertia. Since the coordinates x_1, x_2, x_3 are real, and x_4 is purely imaginary in every system of coordinates,

$$\begin{aligned} &\alpha_{\iota\kappa} \text{ and } \alpha_{44} \text{ are real,} \\ \text{while} \quad &\alpha_{\iota 4} \text{ and } \alpha_{4\kappa} \text{ are purely imaginary,} \end{aligned} \tag{5}$$

if ι and κ denote arbitrary index values between 1 and 3.

If the variables (x_i) and (x_i') were real numbers, they could be interpreted as Cartesian coordinates of a point in a four-dimensional Euclidean space. Then the transformation (3) would represent a simple rotation of the Cartesian system of coordinates in this space, the distance

(4) from the observed point (x_ι) to the origin $(0,0,0,0)$ being invariant by such a rotation.

Now it is natural, even if x_4 and x'_4 are not real, to introduce a four-dimensional space whose points are *defined* by the coordinates (x_ι). Since any event in physical space is characterized by a definite set of numbers (x_ι) in every system of space-time coordinates, each event is thus depicted in a definite point of this abstract four-dimensional space. This space, which was first introduced by Poincaré† and Minkowski,‡ is called the space-time continuum or simply the four-space or $(3+1)$-space, hereby recalling that the four dimensions of the space are not completely equivalent. A homogeneous Lorentz transformation (3) can thus be interpreted as a rotation of the system of coordinates in the $(3+1)$-space. The invariant form (4) is naturally called the square of the four-dimensional *distance* between the event point (x_i) and the origin $(0,0,0,0)$. In view of the formal similarity to a Euclidean space, all usual geometrical notions can now be used in the $(3+1)$-space. The geometry in this space is called pseudo-Euclidean. The deviation from Euclidean geometry is characterized by the circumstance that the distance (4) can be zero without all (x_ι) being zero. All points whose distance from the origin is zero form a surface described by the equation

$$s^2 = \sum_\iota x_\iota^2 = x^2+y^2+z^2-c^2t^2 = 0. \tag{6}$$

This surface is called the light cone, since the equation (6) describes the propagation of a spherical light wave starting from the origin

$$x = y = z = 0$$

at the time $t = 0$. The light cone divides the $(3+1)$-space into two invariant separate domains a and b characterized by the inequalities

$$s^2 = x^2+y^2+z^2-c^2t^2 < 0 \tag{7 a}$$

and

$$s^2 = x^2+y^2+z^2-c^2t^2 > 0, \tag{7 b}$$

respectively. For events in the latter domain we can always by a Lorentz transformation introduce a system of space-time coordinates S' in which $t' = 0$, i.e. the event $(x_\iota) = (x,y,z,t)$ is simultaneous with the event $(0,0,0,0)$ in the system S'. Such a transformation is not possible for two events in the domain (7 a).

An arbitrary motion of a material particle can be described by equations of the form

$$x_\iota = x_\iota(x_4) \quad (\iota = 1,2,3), \tag{8}$$

where the $x_\iota(x_4)$ are definite functions of the time variable x_4. These

† H. Poincaré, *Rend. Pal.* **21**, 129 (1906).
‡ H. Minkowski, 'Raum und Zeit', *Phys. ZS.* **10**, 104 (1909).

equations represent a curve in the (3+1)-space which we shall call the time track of the particle. In the case of a uniform motion the functions $x_\iota(x_4)$ are linear and the time track is a straight line. If it goes through the origin, it must evidently lie entirely inside the domain (7 a) since the velocity of the particle is always smaller than the velocity of light. Then we can always introduce a system of coordinates S' such that the x_4'-axis coincides with the time track of the particle. In ordinary physical space (3-space) this means simply that we can always introduce a system of inertia which follows the particle in its motion.

Let us now consider two arbitrary events with coordinates (x_ι) and $(\bar{x}_\iota)$ in S By transition to another system of coordinates S' both sets of coordinates are transformed in the same way, viz. by equations of the form (3) with the same coefficients $\alpha_{\iota k}$. This means that the differences $(x_\iota-\bar{x}_\iota)$ between these coordinates are also transformed according to these equations, so that

$$\sum_\iota (x_\iota-\bar{x}_\iota)^2 = \sum_\iota (x_\iota'-\bar{x}_\iota')^2 \tag{9}$$

must be an invariant. The quantity (9) represents the square of the four-dimensional distance between the event points (x_ι) and $(\bar{x}_\iota)$. Since both $\sum_\iota x_\iota^2$ and $\sum_\iota \bar{x}_\iota^2$ are invariants it follows at once from (9) that

$$\sum_\iota x_\iota \bar{x}_\iota = \sum_\iota x_\iota' \bar{x}_\iota' \tag{10}$$

is also an invariant.

Using the expressions (3) for x_ι' and $\bar{x}_\iota'$ in the right-hand side of (10) we obtain

$$\sum_\iota \Big(\sum_l \alpha_{\iota l} x_l\Big)\Big(\sum_m \alpha_{\iota m}\bar{x}_m\Big) = \sum_{l,m} \Big(\sum_\iota \alpha_{\iota l}\alpha_{\iota m}\Big) x_l \bar{x}_m.$$

According to (10) this expression must be equal to $\sum_l x_l \bar{x}_l$ for all independent values of the variables (x_l) and $(\bar{x}_m)$. This is possible, however, only when the coefficients satisfy the relations

$$\sum_\iota \alpha_{\iota l}\alpha_{\iota m} = \delta_{lm}, \tag{11}$$

where

$$\delta_{lm} = \begin{cases} 0 \text{ for } l \neq m \\ 1 \text{ for } l = m \end{cases} \tag{12}$$

is the well-known Kronecker symbol.

The relations inverse to (3) are simply

$$x_k = \sum_\iota x_\iota' \alpha_{\iota k}, \tag{13}$$

for if we use (3) in the right-hand side of (13) we get by means of (11)

$$\sum_\iota x_\iota'\alpha_{\iota k} = \sum_{\iota,l} \alpha_{\iota l} x_l \alpha_{\iota k} = \sum_l x_l\Big(\sum_\iota \alpha_{\iota l}\alpha_{\iota k}\Big) = \sum_l x_l\,\delta_{lk} = x_k.$$

Using (13) in the left-hand side of (10) we then obtain by the same argument as in the derivation of (11)

$$\sum_i \alpha_{li}\,\alpha_{mi} = \delta_{lm}. \tag{14}$$

The conditions (11) and (14) for the coefficients α_{ik}, the so-called orthogonality relations, merely express the fact that the homogeneous Lorentz transformations represent 'rotations' of the system of coordinates in (3+1)-space.

For the determinant formed by the scheme of coefficients α_{ik}, i.e.

$$\alpha = |\alpha_{ik}| = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{vmatrix}, \tag{15}$$

we obtain, using (14) and the multiplication rule for determinants,

$$\alpha^2 = |\alpha_{ik}|^2 = \Big|\sum_l \alpha_{il}\alpha_{kl}\Big| = |\delta_{ik}| = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1,$$

i.e.

$$\alpha = \pm 1. \tag{15'}$$

For proper rotations α must be equal to $+1$, since a rotation may be performed in a continuous way, and the identical transformation $x'_i = x_i$ has a scheme of coefficients whose determinant is 1. However, for reflections where one or three of the axes change their sign, α will have the value -1.

The coefficients (α_{ik}) for the special Lorentz transformation (II. 24) are given by the following scheme

$$\alpha_{ik} = \begin{pmatrix} \gamma & 0 & 0 & iv\gamma/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -iv\gamma/c & 0 & 0 & \gamma \end{pmatrix}, \tag{16}$$

where $\gamma = (1-v^2/c^2)^{-\frac{1}{2}}$. It is easily verified that (16) satisfies the orthogonality relations (11) and (14) and that $\alpha = +1$.

The special Lorentz transformation can also be written in the form

$$\left.\begin{aligned} x'_1 &= x_1\cos\psi + x_4\sin\psi \\ x'_4 &= -x_1\sin\psi + x_4\cos\psi \\ x'_2 &= x_2, \quad x'_3 = x_3 \end{aligned}\right\}, \tag{17}$$

where we have put

$$\cos\psi = \gamma, \qquad \sin\psi = iv\gamma/c, \qquad \tan\psi = iv/c. \tag{18}$$

Formally the equations (17) represent a rotation in the $(x_1 x_4)$-plane, but the 'angle of rotation' ψ is a purely imaginary quantity.

34. Lorentz contraction and retardation of moving clocks in four-dimensional representation

Any event which takes place on the x-axis in the system of inertia S is represented by a point in the $(x_1 x_4)$-plane in (3+1)-space. Fig. 12,

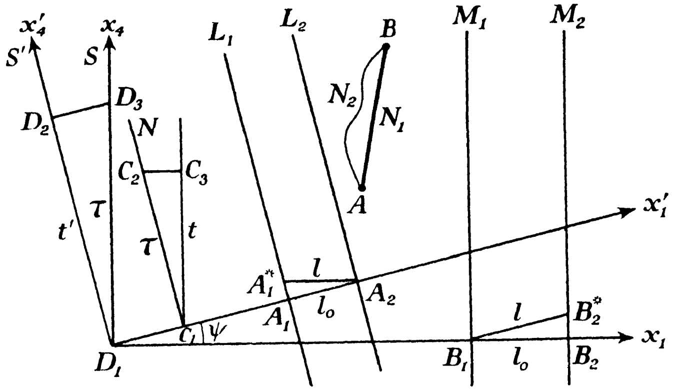

FIG. 12.

which is drawn as if the time variables x_4, x_4', and the angle ψ were real, gives an illustration of the Lorentz contraction and the retardation of clocks. The lines L_1 and L_2, which are parallel to the x_4'-axis, represent the time tracks of the end-points of a measuring-rod at rest on the x'-axis, the rest length being $l_0 = x'_{A_2} - x'_{A_1}$.

The length l of the measuring-rod relative to S will then be equal to the difference in the x-coordinates of two event points A_2 and A_1^* which are simultaneous (and thus have the same x_4-value) in S, i.e.

$$l = x_{A_2} - x_{A_1^*}.$$

Since all formal relations from Euclidean geometry are valid here also, it follows from the figure that

$$l = l_0 \sec\psi = l_0(1-v^2/c^2)^{\frac{1}{2}}, \tag{19}$$

where we have used (18). Equation (19) is the Lorentz contraction formula. Similarly, if we consider a measuring-rod at rest on the x-axis of the system S, the time tracks of the end-points are given by the lines M_1 and M_2 parallel to the x_4-axis. A consideration of the triangle $B_1 B_2 B_2^*$ thus leads again to the equation (19).

In the same way, we may get a geometrical illustration of the clock retardation effect. Consider a clock at rest in the system S'; its time track is given by a straight line parallel to the x'_4-axis, and the proper time τ is a measure of the length of this line. Now let N be this line, and let $C_1 C_2$ be that part of the line which corresponds to the proper time τ. The corresponding time t in S is obtained as the projection of $C_1 C_2$ on the x_4-axis. Then a consideration of the triangle $C_1 C_2 C_3$ gives immediately

$$t = \tau \cos\psi = \tau\gamma, \tag{20}$$

in accordance with formula (II. 36) for the clock retardation.

On the other hand, if we have a clock which is at rest in the system S, a consideration of the triangle $D_1 D_2 D_3$ gives again formula (20) for the connexion between the time t' of the system S' and the proper time of the clock.

In this representation the Lorentz contraction and the clock retardation effect appear as a kind of perspective shortening of measuring-rods and time intervals. However, it must be emphasized that such a representation is quite formal; this is also manifest from the fact that the 'angle of projection' ψ is an imaginary quantity. On the whole, the epistemological significance of the four-dimensional representation should not be exaggerated. In spite of the formal symmetry between the description of space and time in the theory of relativity there is still a fundamental physical difference between the space and time variables. This difference is intimately connected with the difference of the measuring instruments, clocks and measuring-rods, which are required for the physical definition of these variables. Therefore it is not possible by any admissible 'rotation' (3) satisfying the conditions (5) to transform the time axis into a space axis Only those rotations have a physical meaning in which the x_4-axis remains within the domain (7 a), i.e. inside the light cone (6).

35. Covariance of the laws of nature in four-dimensional formulation

In spite of its purely formal character, the four-dimensional representation has been of great significance in the development of the theory of relativity, since it allows us to express the covariance of the laws of nature under Lorentz transformations in a particularly simple way. Each law of nature expresses a certain relation between physical quantities. These quantities are defined by the procedure to be applied in their measurement. Let A, B,... be such a set of physical quantities measured by physicists in a certain system of inertia S. Among the quantities A, B,... some may be so-called field quantities which are functions of the space-time coordinates (x_i) in the system S. Thus a law of nature can be expressed by one or several equations of the form

$$F\left(A, B,..., \frac{\partial A}{\partial x_i}, \frac{\partial B}{\partial x_k},...\right) = 0, \tag{21}$$

where F is a function of the quantities $A, B,...$ and possibly of their derivatives of arbitrarily high order with respect to the space-time coordinates.

Physicists in another system of inertia S' will now by means of the same measuring methods generally find other values $A', B',...$ for the physical quantities mentioned, and if, for example, A is a field variable, A' will generally be a different function of the coordinates (x'_i) than A is of (x_i). Thus the field variables as functions of the space-time coordinates are generally not form-invariant.

However, the physical law expressed by (21) in S can in S' be expressed by equations of the form

$$F\left(A', B',...; \frac{\partial A'}{\partial x'_i}, \frac{\partial B'}{\partial x'_k}, ..\right) = 0, \tag{21'}$$

where the function F in (21′) on account of the special principle of relativity must be the *same* function of the arguments $(A', B',...)$ as is the function F of $(A, B,...)$ in (21), i.e. any relation between physical quantities must be expressed by means of form-invariant or covariant equations.

When the problem arose of expressing the fundamental equations in the theories of electrodynamics and elasticity in a form independent of the Cartesian system of coordinates applied in the description, the three-dimensional vector and tensor calculus was invented. Since the Lorentz transformations represent rotations in $(3+1)$-space, it is therefore natural to attempt to meet the requirement of covariance of the laws of nature under these transformations by a generalization of the three-dimensional vectors and tensors to four dimensions and to write the fundamental equations in the form of four-dimensional tensor equations.

As we shall see in the following chapters, this was possible for all fundamental equations in classical macroscopic physics, and for some time it was even believed that *all* laws of nature could be written in tensor form. With Dirac's quantum mechanical theory of the electron† it became clear, however, that it is necessary for certain physical systems to deal with other quantities besides tensors, the so-called spinors† which

† P. A. M. Dirac, *The Principles of Quantum Mechanics* (3rd ed., Oxford 1947), chap. XI.

have quite different transformation properties, but nevertheless satisfy covariant differential equations of the type of (21), (21′).

36. The four-dimensional line element or interval. Four-vectors

Consider two neighbouring points P and P' in (3+1)-space with the coordinates (x_i) and (x_i+dx_i) in an arbitrary system of coordinates S. On account of (9) the square of the four-dimensional distance between these points is given by

$$ds^2 = \sum_i dx_i^2. \tag{22}$$

This expression for the *line element* or the *interval* defines the geometry in (3+1)-space. The infinitesimal line connecting P and P' is now the prototype of a vector, just as in a three-dimensional space. This vector is defined by its four 'components' (dx_i) relative to an arbitrary system of coordinates S. By a rotation of the system of coordinates leading to the system S' these components transform like the coordinates, i.e. we have

$$dx_i' = \sum_k \alpha_{ik}\, dx_k. \tag{23}$$

Now, *a four-vector is quite generally defined as a quantity which relative to every system of coordinates has four components (a_i) which are transformed in the same way as coordinates (x_i).* Thus, to the rotation (3), (13) of the system of coordinates correspond the transformation equations

$$a_i' = \sum_k \alpha_{ik} a_k, \qquad a_k = \sum_i a_i' \alpha_{ik} \tag{24}$$

for the components of a four-vector.

In analogy with (4) it then follows from (11) and (14) that the '*square of the magnitude of the four-vector*', or the *norm* of the vector

$$\sum_i a_i^2 = \sum_i a_i'^2 \tag{25}$$

is an invariant.

According as the invariant (25) is less than zero, equal to zero, or larger than zero, we speak of a *time-like* vector, a *zero* vector, and a *space-like* vector, respectively.

For the infinitesimal vector (dx_i) connecting the two neighbouring events P and P', (25) becomes identical with (22), viz.

$$ds^2 = \sum_i dx_i^2 = d\sigma^2 - c^2\, dt^2, \tag{26}$$

where $d\sigma$ is the spatial distance between the two space points in physical space in which the two events occur, while dt is the difference in the time of occurrence of the events. The light cone with P as centre is defined by the equation $ds^2 = 0$. If P' lies on the light cone, the two events can be connected by means of a light signal. If the vector (dx_ι) is time-like, P' lies inside the light cone, while it lies outside the cone when (dx_ι) is space-like.

From two four-vectors with the components (a_ι) and (b_ι) can be formed a new vector with the components $(a_\iota+b_\iota)$.

Since, furthermore, the invariant square of the magnitude of this vector can be written

$$\sum_\iota (a_\iota+b_\iota)^2 = \sum_\iota a_\iota^2+\sum_\iota b_\iota^2+2\sum_\iota a_\iota b_\iota,$$

it follows immediately that the quantity

$$\sum_\iota a_\iota b_\iota = \sum_\iota a'_\iota b'_\iota \tag{27}$$

is also an invariant in analogy with (10). The quantity (27) is called the *scalar product* of the two vectors. Two vectors are said to be orthogonal to each other when their scalar product is zero.

The first three components of a four-vector behave under spatial rotations like the components of an ordinary spatial vector $\mathbf{a}$. Therefore, in any system of inertia we can split a four-vector into a spatial part and a temporal part

$$(a_\iota) = (\mathbf{a}, a_4), \tag{28}$$

but this splitting is of course not invariant under Lorentz transformations. As the quantities $(\mathbf{a}, a_4)$ transform in the same way as

$$(\mathbf{x}, x_4) = (\mathbf{x}, ict),$$

we have obviously for an arbitrary Lorentz transformation without rotation, according to (II. 25′),

$$\left.\begin{aligned} \mathbf{a} &= \mathbf{a}'+\frac{\mathbf{v}}{v^2}\,\frac{(\mathbf{v}\,.\,\mathbf{a}')\{1-(1-v^2/c^2)^{\frac{1}{2}}\}-iv^2a_4'/c}{(1-v^2/c^2)^{\frac{1}{2}}} \\ a_4 &= \frac{a_4'+i(\mathbf{v}\,.\,\mathbf{a}')/c}{(1-v^2/c^2)^{\frac{1}{2}}} \end{aligned}\right\}. \tag{29}$$

When the vector is time-like, i.e.

$$\sum_\iota a_\iota^2 = |\mathbf{a}|^2-|a_4|^2 < 0, \tag{30a}$$

we can always introduce a system of coordinates S' such that the spatial vector $\mathbf{a}' = 0$ in this system. We have simply to choose the time axis in S' in the direction of the four-vector (a_ι). If, however, the vector is space-like,

$$\sum_\iota a_\iota^2 = |\mathbf{a}|^2-|a_4|^2 > 0, \tag{30b}$$

we can find a system S' in which $a_4' = 0$. This obviously means that the new time axis is perpendicular to the vector (a_i), for, if (b_ι) denotes a vector in the direction of the new time axis, $b'_\iota = 0$ for $\iota = 1, 2, 3$, i.e.

$$a_\iota b_\iota = a'_\iota b'_i = 0.$$

37. Four-velocity and acceleration. Wave-number vector. Four-ray velocity

Let us now consider the motion of a material particle and the corresponding time track (8) in the (3+1)-space. We can also use a parameter representation for the time track

$$x_\iota = x_\iota(s) \quad (i = 1, 2, 3, 4) \tag{31}$$

with the length s of the curve as parameter. Two neighbouring event points on this curve are connected by an infinitesimal four-vector (dx_ι) with the length ds given by (22) or (26). Since the velocity of a particle is always less than c, it follows from (26) that ds^2 is negative everywhere on the curve. It is therefore convenient to introduce instead of s a new real parameter τ defined by

$$s = ic\tau. \tag{32}$$

Using (32) in (26) we obtain

$$-c^2\, d\tau^2 = d\sigma^2 - c^2\, dt^2, \tag{33}$$

and, since the velocity of the particle u is equal to $d\sigma/dt$, we get

$$d\tau = (1-u^2/c^2)^{\frac{1}{2}}\, dt. \tag{34}$$

Thus τ is identical with the proper time of the particle, i.e. with the time measured by a standard clock which follows the motion of the particle. The proper time of a particle is a measure of the length of the time track.

The time track corresponding to a uniform straight motion is represented by a straight line as, for example, the line N_1 in Fig. 12, which connects the two events A and B. The length of this line is given by

$$\frac{s_1}{ic} = \tau_1 = (t_B - t_A)\sqrt{\left(1 - \frac{u_1^2}{c^2}\right)}, \tag{35}$$

where u_1 is the velocity in the uniform motion and $t_B - t_A$ is the difference in time between the two events A and B. If we consider another arbitrary time track N_2, connecting the same events A and B, it obviously represents a non-uniform motion. The length of this curve is then, according to (34), given by

$$\frac{s_2}{ic} = \tau_2 = \int\limits_{N_2} \left(1 - \frac{u_2^2}{c^2}\right)^{\frac{1}{2}} dt. \tag{36}$$

Since the expressions (35) and (36) hold in any system of coordinates we can, for example, perform the calculation in a system of coordinates whose time axis is parallel to the straight time track N_1. This means only that we introduce a system of inertia following the uniformly moving particle. In this system of coordinates $u_1 = 0$, while the velocity u in the non-uniform motion corresponding to N_2 cannot possibly be zero along the whole track. Therefore we must have

$$\frac{s_2}{ic} < \frac{s_1}{ic}. \tag{37}$$

The straight motion is thus characterized by the fact that the length of the time track in this case has a *stationary* value, viz. a *maximum*, compared with all other possible motions connecting the same two events.

Since the coordinate increments (dx_i) on the time track of a particle are the components of a four-vector and $d\tau$ is an invariant,

$$U_i = \frac{dx_i}{d\tau} \tag{38}$$

will also be components of a four-vector which is called the *four-velocity*. From (38) and (34) we see that the components of U_i are

$$U_i = \left(\frac{\mathbf{u}}{(1-u^2/c^2)^{\frac{1}{2}}}, \frac{ic}{(1-u^2/c^2)^{\frac{1}{2}}}\right), \tag{39}$$

where

$$\mathbf{u} = \frac{d\mathbf{x}}{dt} \tag{40}$$

is the usual three-dimensional velocity vector. The four-velocity is a four-vector which lies in the direction of the tangent of the time track and the norm of this vector has the constant value $-c^2$, for from (39) one obtains

$$\sum_i U_i^2 = \frac{u^2}{1-u^2/c^2} - \frac{c^2}{1-u^2/c^2} = -c^2. \tag{41}$$

Thus the four-velocity is a time-like vector of constant magnitude and by differentiation of (41) we get

$$\sum_i U_i \frac{dU_i}{d\tau} = 0. \tag{41'}$$

If we apply the equation (29) to the four-vector U_i given by (39) we come back to the transformation equation (II. 55′) for velocities and to the equation inverse to (II. 56).

The four-vector

$$A_i = \frac{dU_i}{d\tau} = \dot{U}_i = (1-u^2/c^2)^{-\frac{1}{2}} \frac{dU_i}{dt}$$

is called the four-acceleration and from (39) we get for its components

$$A_i = \left(\frac{c^2}{c^2-u^2}\mathbf{a}+\mathbf{u}\frac{c^2(\mathbf{u}.\mathbf{a})}{(c^2-u^2)^2}, \frac{ic^3(\mathbf{u}.\mathbf{a})}{(c^2-u^2)^2}\right), \tag{42}$$

where
$$\mathbf{a} = \frac{d\mathbf{u}}{dt} = \frac{d^2\mathbf{x}}{dt^2}$$

is the usual three-dimensional acceleration of the particle. According to (41′) the vectors U_i and A_i are orthogonal to each other. In the rest system S' of the particle the components of A_i are

$$A'_i = (\mathbf{a}', 0). \tag{42'}$$

Another example of a four-vector is given by the wave number vector (σ_i) of a plane monochromatic wave The invariant phase F of the wave in an arbitrary system of coordinates S can, on account of (II. 68), be written

$$-F = \sum_i \sigma_i x_i, \qquad \sigma_i = \left(\frac{\nu}{w}\mathbf{n}, i\frac{\nu}{c}\right) = (\boldsymbol{\sigma}, \sigma_4), \qquad |\boldsymbol{\sigma}| = \frac{\nu}{w} = \frac{1}{\lambda}. \tag{43}$$

Here, $\mathbf{n}$ denotes a unit three-vector in the direction of the wave normal, ν is the frequency w is the phase velocity, and λ is the wave-length of the wave. As the phase is an invariant we have, according to (3),

$$\sum_k \sigma_k x_k = \sum_i \sigma'_i x'_i = \sum_{i,k} \sigma'_i \alpha_{ik} x_k,$$

and, since this equation must hold for arbitrary x_k, we get

$$\sigma_k = \sum_i \sigma'_i \alpha_{ik}, \tag{44}$$

which shows that (σ_i) is a four-vector. Equations (44), or their equivalent (29), yield directly formulae (70), (71), and (72) deduced in § 23 for the transformation of the characteristics of a plane wave.

In § 24, Chapter II, we saw that the ray velocity in a medium with refractive index $n > 1$ transforms like the velocity of a particle with the velocity $u' = w' = c/n$ in the rest system of the medium. In analogy with (39) we can therefore define a four-ray velocity vector (U_i) with the components

$$U_i = \left\{\frac{u.\mathbf{e}}{\sqrt{(1-u^2/c^2)}}, \frac{ic}{\sqrt{(1-u^2/c^2)}}\right\}, \tag{45}$$

where u is the magnitude of the ray velocity and $\mathbf{e}$ is a unit three-vector in the direction of the ray.

In the rest system S' of the refractive medium where $\mathbf{e}' = \mathbf{n}'$ and $w' = u' = c/n$, we have

$$\sigma'_i = \left(\frac{\nu' n}{c}\mathbf{n}', i\frac{\nu'}{c}\right), \qquad U'_i = \left\{\frac{c\mathbf{n}'}{\sqrt{(n^2-1)}}, \frac{icn}{\sqrt{(n^2-1)}}\right\}. \tag{46}$$

In another arbitrary system of inertia S we then have $U_k = \sum_i U'_i \alpha_{ik}$. From (46) it follows that

$$\sum_i \sigma_i U_i = \sum_i \sigma'_i U'_i = 0, \tag{47}$$

i.e. the two four-vectors σ_i and U_i are orthogonal to each other. Inserting (43) and (45) we obtain

$$\frac{\nu}{\sqrt{(1-u^2/c^2)}}\left\{\frac{u}{w}(\mathbf{n}.\mathbf{e})-1\right\} = 0$$

or

$$w = (\mathbf{u}.\mathbf{n}). \tag{47'}$$

(47) thus expresses simply that the phase velocity is equal to the projection of the ray velocity in the direction of the wave normal in any system of inertia.

While the square of the magnitude of (U_i) is constant, equal to $-c^2$, we have

$$\sum_i \sigma_i^2 = \nu^2\left(\frac{1}{w^2}-\frac{1}{c^2}\right) = \frac{\nu'^2}{c^2}(n^2-1). \tag{48}$$

Thus, for a light wave in empty space the phase wave vector is a zero vector.

38. Four-momentum. Four-force. Fundamental equations of point mechanics in four-dimensional vector form

In § 28, formulae (32) and (37), we have seen that momentum and energy ($\mathbf{p}$, E/c) of a material particle transform in the same way as the space-time coordinates ($\mathbf{x}, ct$). Therefore the four quantities

$$p_i = (\mathbf{p}, iE/c) \tag{49}$$

are the components of a four-vector, the four-momentum vector. When we use (III. 23) and (III. 31) in (49) we see that the four-momentum p_i is proportional to the four-velocity (39) and the factor of proportionality is equal to the rest mass m_0 of the particle, i.e.

$$p_i = m_0 U_i. \tag{50}$$

On account of (50), (49), and (41) the norm of this vector is

$$\sum_i p_i^2 = p^2 - \frac{E^2}{c^2} = m_0^2 \sum_i U_i^2 = -m_0^2 c^2 \tag{51}$$

in accordance with (III. 35) The four-momentum vector is thus a time-like vector.

Since p_i in (50) and σ_i in (43) are transformed in the same way, it is possible in an invariant way to adjoin a plane wave with the wave number vector σ_i to a free particle with four-momentum p_i, such that

$$p_i = h\sigma_i, \tag{52}$$

where h is an invariant constant. If h is chosen equal to Planck's constant, the adjoined wave is the de Broglie wave† of the particle. The phase velocity w of the de Broglie wave is, according to (43), (52), and (III. 36), connected with the velocity u of the particle by the equation

$$\frac{1}{w} = \frac{i|\boldsymbol{\sigma}|}{c\sigma_4} = \frac{i|\mathbf{p}|}{cp_4} = \frac{p}{E} = \frac{u}{c^2}. \tag{52'}$$

The theorems of conservation of momentum and energy in a collision between particles can now be comprised in the equation

$$p_\iota = \bar{p}_\iota, \tag{53}$$

where p_ι and $\bar{p}_\iota$ are the sum of the four-momentum vectors of the particles before and after the collision, respectively. The first three equations ($i = 1, 2, 3$) represent the theorem of conservation of momentum while the fourth equation, corresponding to $\iota = 4$, expresses the conservation of energy.

In (53) these theorems are expressed in a form in which the covariance under Lorentz transformations is obvious, for two four-vectors whose components are equal in one system of coordinates will obviously have equal components in any system of coordinates.

As mentioned in the conclusion of § 28, the quantities $(\mathbf{F}_M, (\mathbf{F}_M \cdot \mathbf{u}))$ transform in the same way as $(\mathbf{p}, E)$, i.e. the four quantities

$$F_\iota = \{\mathbf{F}_M, i(\mathbf{F}_M \cdot \mathbf{u})/c\} = \left\{\frac{\mathbf{F}}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{F}\cdot\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right\} \tag{54}$$

are the components of a four-vector, Minkowski's four-force.‡ The fundamental equations of mechanics (III. 42) can thus be written in the following covariant four-vector form

$$\frac{dp_\iota}{d\tau} = F_\iota, \tag{55}$$

or, since m_0 is constant in time,

$$m_0 \frac{dU_\iota}{d\tau} = F_\iota. \tag{56}$$

The first three equations are the equations of motion, and the fourth equation expresses the theorem of conservation of energy. Conversely, from the validity of (55) in every system of inertia it follows at once that F_ι is a four-vector, for F_ι must then transform in the same way as the left-hand side of (55) which is obviously a four-vector, since τ is invariant.

† Op cit, chap. II, p. 58

‡ H. Minkowski, 'Das Relativitatsprinzip', *Ann. d. Phys.* **47**, 927 (1915).

From (41′) and (56) it is seen that the vectors U_ι and F_ι are orthogonal to each other, viz.

$$\sum_\iota F_\iota U_\iota = 0. \tag{57}$$

This property of F_ι is closely connected with the constancy in time of the rest mass. From (55), (41), and (41′) we obtain

$$\sum_\iota F_\iota U_\iota = \sum_\iota U_\iota \frac{d(m_0 U_\iota)}{d\tau} = \sum_\iota U_\iota^2 \frac{dm_0}{d\tau} + m_0 \sum_\iota U_\iota \frac{dU_\iota}{d\tau} = -c^2 \frac{dm_0}{d\tau}. \tag{58}$$

Equation (57) thus expresses the fact that the proper mass m_0 is conserved.

Sometimes we are concerned with systems in which the external forces produce a change in the proper mass of the particle. This is the case, for instance, in an electrically conducting substance under the influence of electromagnetic forces, since the Joule heat energy produced in the body will contribute to the proper mass. If in such a case we wish to maintain the equations (55), we must in the expression (54) for F_ι replace $\mathbf{F}.\mathbf{u} = A$ by the total effect. While F_1, F_2, F_3 remain unchanged, F_4 will be defined by

$$F_4 = \frac{\iota(A+Q)}{\sqrt{(c^2-u^2)}}, \tag{59}$$

where

$$A = (\mathbf{F}.\mathbf{u}) \tag{60}$$

is the mechanical work performed per unit of time, while Q is the amount of heat or non-mechanical energy developed per unit of time in the body. The fourth equation (55) which, by means of (49), (59), and (34) can be written

$$\frac{dE}{dt} = A+Q, \tag{61}$$

thus again expresses the conservation of energy. As the equations (55) must hold in every system of inertia, and since the left-hand side is a four-vector, the generalized four-force must also be a four-vector. But now (F_ι) is no longer orthogonal to (U_ι). Instead, we have

$$\sum_\iota F_\iota U_\iota = \frac{(\mathbf{F}.\mathbf{u})}{1-u^2/c^2} - \frac{A+Q}{1-u^2/c^2} = -\frac{Q}{1-u^2/c^2}. \tag{62}$$

Since the left-hand side of this equation is an invariant,

$$\frac{Q}{1-u^2/c^2} = \frac{Q'}{1-u'^2/c^2} = Q^0 \tag{63}$$

must thus be an invariant. It is equal to the amount of heat developed in the rest system per unit time.

In a physical system of this kind the rest mass of the particle is no longer conserved. From (58), (62), and (63) we get

$$\frac{dm_0}{d\tau} = \frac{Q^0}{c^2}. \tag{64}$$

Since τ is identical with the time in the rest system, (64) simply means that the amount of heat developed in the rest system has an inertial mass corresponding to Einstein's equation (III. 74).

While the equation (55) thus holds quite generally, (56) is valid only when the rest mass is conserved, i.e. when $\sum F_\iota U_\iota = 0$.

The equation (55) may also be written

$$m_0 \frac{dU_\iota}{d\tau} + \frac{dm_0}{d\tau} U_\iota = F_\iota,$$

or, on account of (58),

$$m_0 \frac{dU_\iota}{d\tau} = F_\iota + \frac{\sum F_k U_k}{c^2} U_\iota. \tag{55'}$$

Thus, when $\sum F_\iota U_\iota \neq 0$, i.e. *when the proper mass is not conserved, a force will in general be necessary in order to maintain a uniform motion* of the particle; for $dU_\iota/d\tau = 0$ only if

$$F_\iota = -\frac{\sum (F_k U_k)}{c^2} U_\iota = \frac{Q^0}{c^2} U_\iota. \tag{65}$$

In the rest system this means $\mathbf{F}^0 = 0$, but in every other system of inertia we have

$$\mathbf{F} = \frac{Q^0}{c^2} \mathbf{u} \neq 0.$$

From (63) we get at the same time the transformation properties of the amount of heat Q conveyed to a system in a process of that kind. Multiplying (63) by $\Delta\tau\sqrt{(1-u^2/c^2)}$ we obtain, using (34),

$$Q\,\Delta t = Q^0\,\Delta\tau\sqrt{(1-u^2/c^2)}.$$

$\Delta Q = Q\,\Delta t$ is the total amount of heat conveyed to the system during an interval Δt, while $\Delta Q^0 = Q^0\,\Delta\tau$ is the corresponding quantity measured in the rest system. Thus we have

$$\Delta Q = \Delta Q^0 \sqrt{\left(1-\frac{u^2}{c^2}\right)}. \tag{66}$$

In Chapter VII we shall see that the equation (66) also holds for any amount of heat conveyed to a system in a thermodynamical process.

39. Tensors of rank 2

In the preceding chapters we have seen that the covariance of the fundamental equations of mechanics under Lorentz transformations can be expressed in an especially elegant way by writing them in four-dimensional vector form. To obtain a similar geometrical representation of electrodynamics, for instance, it is necessary to introduce four-dimensional tensors also.

By a tensor of second rank in (3+1)-space we mean a quantity which has 4^2 components (t_{ik}) relative to an arbitrary system of coordinates S and whose components (t'_{ik}) in another arbitrary system S' are connected with the components (t_{ik}) by means of the equations

$$t'_{ik} = \sum_{l,m} \alpha_{il}\alpha_{km}t_{lm} = \alpha_{il}\alpha_{km}t_{lm}, \qquad t_{ik} = t'_{lm}\alpha_{li}\alpha_{mk}, \tag{67}$$

where the α_{ik} are the same coefficients as in equation (3) which defines the transition from S to S'.

For the sake of simplicity we have here omitted the sign of summation, substituting it in the following by the convention that an expression in which a Latin index, like l or m in (67), appears twice shall be summed over this index from 1 to 4. Free indices like i and k in (67) can assume independently the values 1, 2, 3, 4. Equation (25), for example, will thus in the future be written $a_i a_i = a'_i a'_i$. If an index can assume only the values 1, 2, 3 it will be denoted by a Greek letter, and if it appears twice in an expression it is implied that we shall sum over this index from 1 to 3. The square of the magnitude of a spatial vector will thus be written

$$|\mathbf{a}|^2 = a_\iota a_\iota.$$

The definition (67) of a four-tensor is a direct generalization of an ordinary spatial tensor whose components by rotations in physical space transform according to the equations

$$t'_{\iota\kappa} = \alpha_{\iota\lambda}\alpha_{\kappa\mu}t_{\lambda\mu}, \tag{68}$$

where $\alpha_{\iota\kappa}$ are the coefficients in the orthogonal transformation representing the rotation.

The sum of the diagonal elements of a tensor of second rank is an invariant, for from (67) and (11) we obtain

$$t'_{ii} = (\alpha_{il}\alpha_{im})t_{lm} = \delta_{lm}t_{lm} = t_{ll}. \tag{69}$$

In the same way it is seen by means of the orthogonality relations (11), (14) that the quantity

$$t_{ik}t_{ik} = t'_{lm}t'_{lm} \tag{70}$$

is an invariant.

If α_{ik} represents a simple rotation in the three-dimensional physical space, we have

$$\alpha_{\iota 4} = \alpha_{4\kappa} = 0, \qquad \alpha_{44} = 1, \tag{71}$$

and (67) is reduced to

$$\left.\begin{aligned} t'_{\iota\kappa} &= \alpha_{\iota\lambda}\alpha_{\kappa\mu}t_{\lambda\mu}, & t'_{\iota 4} &= \alpha_{\iota\lambda}t_{\lambda 4} \\ t'_{4\kappa} &= \alpha_{\kappa\mu}t_{4\mu}, & t'_{44} &= t_{44} \end{aligned}\right\}. \tag{72}$$

(72) shows that the spatial part of a four-tensor by spatial rotations behaves like an ordinary three-dimensional tensor. Furthermore, the numbers $t_{\iota 4}$ and $t_{4\kappa}$ separately form the components of a spatial vector, while t_{44} is an invariant for purely spatial rotations.

From a vector a_ι and a tensor $t_{\iota k}$ a new vector can be formed whose components in every system of coordinates are given by the equation

$$b_\iota = t_{\iota k}a_k, \tag{73}$$

for from (24), (67), and (11) we obtain

$$\begin{aligned} b'_\iota &= t'_{\iota k}a'_k = \alpha_{\iota l}\alpha_{km}t_{lm}\alpha_{kn}a_n = \alpha_{\iota l}(\alpha_{km}\alpha_{kn})t_{lm}a_n \\ &= \alpha_{\iota l}\,\delta_{mn}t_{lm}a_n = \alpha_{\iota l}(t_{lm}a_m) = \alpha_{\iota l}b_l. \end{aligned} \tag{74}$$

If $a_{\iota k}$ and $b_{\iota k}$ are the components of two tensors, $a_{\iota k}+b_{\iota k}$ will obviously also be the components of a tensor which is called *the sum of the two tensors*. Further, if $t_{\iota k}$ is a tensor, $\check{t}_{\iota k} = t_{k\iota}$ is also a tensor, viz. the transposed tensor. When a_ι and b_ι denote the components of two vectors which thus are transformed according to (24), the quantities

$$t_{\iota k} = a_\iota b_k \tag{75}$$

will obviously be transformed according to (67). The tensor of second rank, defined by (75), is called *the direct product of the vectors* a_ι *and* b_k. Also the quantities

$$t_{\iota k} = a_\iota b_k - a_k b_\iota = -t_{k\iota}, \tag{76}$$

which denote the difference between the tensor (75) and its transposed, thus represent a tensor. A tensor like (76), satisfying the equation

$$t_{\iota k} = -t_{k\iota} = -\check{t}_{\iota k} \tag{77}$$

for all values of the indices i and k, is called antisymmetrical. Analogously, a tensor satisfying the equations

$$t_{\iota k} = t_{k\iota} = \check{t}_{\iota k} \tag{78}$$

is called symmetrical.

Since both sides of the equations (77) and (78) transform like tensors, these equations must hold in every system of coordinates if they hold in one system. Symmetry or antisymmetry of a tensor is thus an invariant property. In an antisymmetrical tensor all diagonal elements are zero, for if we put $i = k$ in (77) (without summing), we obtain

$$t_{ii} = -t_{ii} = 0 \tag{79}$$

for any i.

An antisymmetrical four-tensor of second order $F_{ik} = -F_{ki}$ has only six independent components. Putting

$$H_{\iota\kappa} = F_{\iota\kappa}, \qquad E_\iota = iF_{\iota 4} = -iF_{4\iota}, \tag{80}$$

$H_{\iota\kappa}$ and E_ι will, according to (72), in spatial rotations (71) behave like components of an antisymmetrical spatial tensor and an ordinary space vector $\mathbf{E}$, respectively.

Moreover, putting

$$H_1 = H_{23}, \quad H_2 = H_{31}, \quad H_3 = H_{12}, \tag{80'}$$

we obtain the following transformation equation for H_ι and E_ι in the case of a special Lorentz transformation (16),

$$\left.\begin{array}{lll} H'_1 = H_1, & H'_2 = \gamma(H_2+vE_3/c), & H'_3 = \gamma(H_3-vE_2/c) \\ E'_1 = E_1, & E'_2 = \gamma(E_2-vH_3/c), & E'_3 = \gamma(E_3+vH_2/c) \end{array}\right\}, \tag{81}$$

where $\gamma = (1-v^2/c^2)^{-\frac{1}{2}}$. The corresponding reciprocal equations are obtained by interchanging the primed and unprimed quantities and replacing v by $-v$. With $\mathbf{v} = (v, 0, 0)$ the latter equations may be comprised in the vector formulae

$$\left.\begin{array}{l} \mathbf{E} = \gamma\mathbf{E}' + \dfrac{\mathbf{v}}{v^2}(\mathbf{v}.\mathbf{E}')(1-\gamma) - \dfrac{\gamma}{c}(\mathbf{v}\times\mathbf{H}') \\ \mathbf{H} = \gamma\mathbf{H}' + \dfrac{\mathbf{v}}{v^2}(\mathbf{v}.\mathbf{H}')(1-\gamma) + \dfrac{\gamma}{c}(\mathbf{v}\times\mathbf{E}') \end{array}\right\}, \tag{81'}$$

and in this form they are valid for any Lorentz transformation without rotation of the spatial axes.

40. Angular momentum and moment of force in four-dimensional representation

Let $(x_\iota) = (\mathbf{x}, ict)$ be the space-time coordinates of an event point on the time track of a material particle, and let $(p_i) = \{\mathbf{p}, i(E/c)\}$ be the four-momentum of the particle. According to (76), from these two four-vectors we can form an antisymmetrical tensor

$$M_{\iota k} = x_\iota p_k - x_k p_\iota. \tag{82}$$

The spatial part of this tensor is an antisymmetrical space tensor, the angular momentum tensor $M_{\iota\kappa}$, whose components are connected with the angular momentum vector

$$\mathbf{M} = \mathbf{x} \times \mathbf{p} \tag{83}$$

by the equations

$$\mathbf{M} = (M_x, M_y, M_z) = (M_{23}, M_{31}, M_{12}). \tag{84}$$

In the same way we can from (x_ι) and from the Minkowski four-force (F_ι) form the tensor

$$D_{\iota k} = x_\iota F_k - x_k F_\iota. \tag{85}$$

The spatial part of this tensor gives, analogously to (83), (84), the moment of the Minkowski force (III. 41) relative to the origin.

By means of the fundamental equations of mechanics (55) we obtain, using (38),

$$\frac{dM_{\iota k}}{d\tau} = U_\iota p_k + x_\iota F_k - U_k p_\iota - x_k F_\iota,$$

or, by means of (50) and (85),

$$\frac{dM_{\iota k}}{d\tau} = D_{\iota k}. \tag{86}$$

The spatial part of this equation contains the angular momentum theorem

$$\frac{d\mathbf{M}}{dt} = (\mathbf{x} \times \mathbf{F}). \tag{86'}$$

The Kronecker symbol defined by (12) represents a tensor of second rank of especially simple character. Consider a tensor whose components in S are equal to $\delta_{\iota k}$; on account of (67) and (14) its components in S' are then

$$\delta'_{\iota k} = \alpha_{\iota l}\alpha_{km}\,\delta_{lm} = \alpha_{\iota l}\alpha_{kl} = \delta_{\iota k}. \tag{87}$$

This tensor has thus the same constant components in every system of coordinates and it is the only tensor whose components remain unchanged by a transition to another system of coordinates.

41. Tensors of arbitrary rank

In analogy with (67) a tensor of third rank in (3+1)-space is now defined as a quantity with 4^3 components $t_{\iota kl}$ which transform according to the equations

$$t'_{\iota kl} = \alpha_{\iota m}\alpha_{kn}\alpha_{lp}t_{mnp}, \qquad t_{\iota kl} = t'_{mnp}\alpha_{m\iota}\alpha_{nk}\alpha_{pl}. \tag{88}$$

Thus every index transforms separately according to the same law as for a *four-vector. A four-vector may in fact be regarded as a tensor of first rank. In the same sense an invariant is a tensor of zero rank.*

A tensor of rank n is then a quantity $t_{\iota kl\ldots}$ with n independent indices, each of which transforms separately according to the law (24), characteristic of a vector. If in a tensor of rank n two indices, for example k and i, are put equal to each other, we obtain after summing over this index a tensor $t_{iilm..}$ of rank $(n-2)$. This is a direct consequence of the transformation equations for tensors in connexion with the orthogonality relations (11), (14). Such a process by which from a tensor of rank n a tensor of rank $(n-2)$ is formed is called *contraction*. Equation (69), in which a tensor of rank zero has been formed from a tensor of rank 2, represents a special case of such a contraction.

By addition (or subtraction) of corresponding components of two tensors of rank n we naturally obtain a new tensor of rank n. On the other hand, it has no covariant meaning to add two tensors of different rank. However, we can always form *the direct product* of two tensors of ranks n and m, respectively, by forming all possible products of the components of these tensors. Hereby we obtain a new tensor of rank $(n+m)$. The equation (75) obviously represents a special case of this general theorem, since the tensor $t_{ik} = a_i b_k$ of rank 2 is the direct product of the two tensors of first rank a_i and b_k. By subsequent contraction of the tensor $a_i b_k$ we obtain a tensor of rank zero, viz. the invariant (27).

Equation (73) also represents a special case of a combination of both operations: direct multiplication and contraction. Primarily, a tensor $(t_{ik}.a_l)$ of rank 3 is formed by direct multiplication of the two tensors t_{ik} and a_l. By contraction we then obtain the tensor $b_i = t_{ik} a_k$ of first rank. In the same way, equation (70) can be regarded as a result of a direct multiplication of t_{ik} by itself succeeded by two contractions.

42. Pseudo-tensors

In three-dimensional vector calculus one introduces besides the ordinary (polar) vectors with the transformation law

$$a'_\iota = \alpha_{\iota\kappa} a_\kappa \tag{89}$$

so-called axial vectors which transform according to the equations

$$a'_\iota = \alpha\alpha_{\iota\kappa} a_\kappa, \tag{90}$$

where $\alpha = |\alpha_{\iota\kappa}|$ is the transformation determinant. For proper rotations we have $\alpha = 1$, and an axial vector transforms like a polar vector. However, by reflections in which one or three axes change their signs we have $\alpha = -1$: thus, for instance, by a reflection at the origin in which

$$x'_\iota = -x_\iota, \tag{91}$$

the components of an axial vector are unchanged, while the components of a polar vector change their signs. An example of an axial vector is the vector product **c** of two polar vectors **a** and **b** with the components

$$\mathbf{c} = (c_1, c_2, c_3) = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1). \tag{92}$$

By a reflection $a'_\iota = -a_\iota$, $b'_\iota = -b_\iota$, we have obviously $c'_\iota = c_\iota$. Another well-known example of an axial vector is the magnetic field vector **H**.

The generalization of the notion of an axial vector to tensors of higher rank and to four dimensions is obvious. These quantities are called *pseudo-tensors*. They transform like tensors, except that they are also multiplied by the transformation determinant $\alpha = |\alpha_{ik}|$ defined by (15). Thus, a pseudo-tensor of rank 2 is a quantity with 4^2 components in every system of coordinates with the transformation law

$$t'_{ik} = \alpha\alpha_{il}\alpha_{km}t_{lm}. \tag{93}$$

From this definition we get at once the following rules. The sum of two pseudo-tensors of the same rank is again a pseudo-tensor of equal rank. The direct product of a pseudo-tensor and a tensor is a pseudo-tensor with a rank equal to the sum of the ranks of the two factors in the product. The direct product of two pseudo-tensors is a tensor. The operation of contraction can be performed with pseudo-tensors in the same way as with tensors, thus leading to a pseudo-tensor whose rank is diminished by 2.

43. The Levi-Civita symbol

Like the Kronecker symbol, which was shown in § 40 to be a tensor with constant components in every coordinate system, the Levi-Civita symbol is a pseudo-tensor with the same property. In four-dimensional space this pseudo-tensor is of rank 4. *The Levi-Civita symbol is defined as a quantity δ_{iklm} which is antisymmetric in all four indices.* Thus, *the only non-vanishing components of δ_{iklm} are those for which all four indices are different and they are equal to $+1$ or -1 according as (i, k, l, m) is an even or an odd permutation of* $(1, 2, 3, 4)$. Now consider a pseudo-tensor which in S has the components δ_{iklm}. In another system S' its components are then

$$\delta'_{iklm} = \alpha\alpha_{in}\alpha_{kp}\alpha_{lq}\alpha_{mr}\delta_{npqr}. \tag{94}$$

Since all symmetry properties are conserved by the transformation, δ'_{iklm} is also antisymmetric in all indices and we need only calculate the component with $(i, k, l, m) = (1, 2, 3, 4)$ for which we get

$$\delta'_{1234} = \alpha\alpha_{1n}\alpha_{2p}\alpha_{3q}\alpha_{4r}\delta_{npqr}. \tag{95}$$

From the definition of δ_{iklm} it follows that

$$\alpha_{1n}\,\alpha_{2p}\,\alpha_{3q}\,\alpha_{4r}\,\delta_{npqr} = \alpha,$$

thus, by means of (95) and (15′),

$$\delta'_{1234} = \alpha^2 = 1 = \delta_{1234}.$$

From the symmetry properties of δ'_{iklm} and δ_{iklm} it then follows that

$$\delta'_{iklm} = \delta_{iklm} \tag{96}$$

for all values of the indices (i, k, l, m), which shows that the Levi–Civita symbol is a pseudo-tensor with the same constant components in every coordinate system.

In three-dimensional space, the Levi–Civita symbol is a quantity $\delta_{\iota\kappa\lambda}$ antisymmetric in all three indices. δ_{123} is equal to 1, the other non-vanishing components following from this by the symmetry rules. It is shown in the same way as before that $\delta_{\iota\kappa\lambda}$ is a three-dimensional pseudo-tensor.

44. Dual tensors

By means of the Levi–Civita symbol we can associate an antisymmetric three-tensor $H_{\iota\kappa}$ with a pseudo-vector (axial vector) $\mathbf{H}$ by

$$H_\iota = \tfrac{1}{2}\delta_{\iota\kappa\lambda}\,H_{\kappa\lambda}, \tag{97}$$

i.e.

$$\mathbf{H} = (H_1, H_2, H_3) = (H_{23}, H_{31}, H_{12}). \tag{98}$$

The quantities H_ι defined by (80′) are thus the components of an axial vector dual to the tensor $H_{\iota\kappa}$.

If $H_{\iota\kappa}$ is of the form

$$\sigma_{\iota\kappa} = a_\iota b_\kappa - a_\kappa b_\iota = \begin{vmatrix} a_\iota & b_\iota \\ a_\kappa & b_\kappa \end{vmatrix}, \tag{99}$$

where a_ι and b_κ are two vectors, the corresponding axial vector

$$\sigma_\iota = \delta_{\iota\kappa\lambda}\,a_\kappa\,b_\lambda \tag{100}$$

is the vector product $\boldsymbol{\sigma} = \mathbf{a}\times\mathbf{b}$. This tensor or its dual axial vector represents the parallelogram formed by the vectors $\mathbf{a}$ and $\mathbf{b}$, their components being equal to the projections of the parallelogram on the three coordinate planes. The area σ of the parallelogram is given by the equation

$$\sigma^2 = \sigma_\iota\,\sigma_\iota = \tfrac{1}{2}\sigma_{\iota\kappa}\,\sigma_{\iota\kappa}. \tag{101}$$

The pseudo-vector σ_ι is perpendicular to the parallelogram, for from (100) we get

$$\sigma_\iota a_\iota = a_\iota\,\delta_{\iota\kappa\lambda}\,a_\kappa\,b_\lambda = 0 \quad \text{and} \quad \sigma_\iota b_\iota = 0 \tag{102}$$

on account of the antisymmetry of the Levi–Civita symbol. The vector dual to the tensor

$$H_{\iota\kappa} = \frac{\partial a_\kappa}{\partial x_\iota} - \frac{\partial a_\iota}{\partial x_\kappa}$$

is similarly the axial vector curl **a**.

Three vectors **a**, **b**, and **c** define a parallelepiped which by analogy with (99) is represented by an antisymmetrical tensor

$$V_{\iota\kappa\lambda} = \begin{vmatrix} a_\iota & b_\iota & c_\iota \\ a_\kappa & b_\kappa & c_\kappa \\ a_\lambda & b_\lambda & c_\lambda \end{vmatrix}. \tag{103}$$

By means of the Levi–Civita symbol we can associate this tensor with a pseudo-tensor of rank zero, i.e. with a pseudo-invariant

$$V = \frac{1}{3!}\,\delta_{\iota\kappa\lambda} V_{\iota\kappa\lambda} = \delta_{\iota\kappa\lambda} a_\iota b_\kappa c_\lambda = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \tag{104}$$

V represents the volume of the parallelepiped. It is invariant for proper rotations, but changes sign by reflections. We obviously have

$$V^2 = \frac{1}{3!} V_{\iota\kappa\lambda} V_{\iota\kappa\lambda}. \tag{105}$$

If **a**, **b**, **c** are infinitesimal vectors lying in the directions of the x_1-, x_2-, and x_3-axes, respectively, for instance

$$\mathbf{a} = (dx_1, 0, 0), \qquad \mathbf{b} = (0, dx_2, 0), \qquad \mathbf{c} = (0, 0, dx_3), \tag{106}$$

the corresponding volume element is, by (104) and (106),

$$dV = dx_1\,dx_2\,dx_3. \tag{107}$$

It is a pseudo-invariant.

In (3+1)-space we can associate an antisymmetrical tensor of rank $n \leqslant 4$ with a pseudo-tensor of rank $(4-n)$ by means of the Levi–Civita symbol δ_{iklm}. Thus the dual pseudo-tensor F^*_{ik} to an antisymmetrical tensor F_{ik} is defined by the equations

$$F^*_{ik} = \frac{1}{2i}\,\delta_{iklm} F_{lm}, \tag{108}$$

i.e.

$$\left.\begin{aligned} F^*_{23} &= \frac{1}{i} F_{14}, & F^*_{31} &= \frac{1}{i} F_{24}, & F^*_{12} &= \frac{1}{i} F_{34} \\ F^*_{14} &= \frac{1}{i} F_{23}, & F^*_{24} &= \frac{1}{i} F_{31}, & F^*_{34} &= \frac{1}{i} F_{12} \end{aligned}\right\}. \tag{109}$$

Introducing the quantities **E** and **H** defined by (80) and (80′) we see that the dual pseudo-tensor F^*_{ik} is obtained from F_{ik} by the substitution

E → **H**, **H** → −**E**. The equations (81) and (81′) are unchanged by this substitution, in accordance with the fact that the Lorentz transformation without rotation is a transformation with $\alpha = 1$, so that F^*_{ik} in this case transforms in the same way as F_{ik}.

A tensor σ_{ik} of the special form

$$\sigma_{ik} = a_i b_k - a_k b_i, \tag{110}$$

where a_i and b_k are two vectors, represents the two-dimensional 'parallelogram' defined by the vectors a_i and b_i by analogy with (99). The dual tensor

$$\sigma^*_{ik} = \frac{1}{i}\,\delta_{iklm}\,a_l b_m$$

is orthogonal to the vectors a_i and b_k and to the tensor σ_{ik}.

$$\sigma^*_{ik}\,a_k = \sigma^*_{ik}\,b_k = \sigma^*_{ik}\,\sigma_{ik} = 0. \tag{111}$$

The area σ of the parallelogram is defined by

$$\sigma^2 = \tfrac{1}{2}\sigma_{ik}\,\sigma_{ik} \tag{112}$$

by analogy with (101).

Likewise, to an antisymmetrical tensor of rank 3 corresponds a dual pseudo-tensor of rank 1, i e. a pseudo-vector. If the tensor is of the form

$$V_{ikl} = \begin{vmatrix} a_i & b_i & c_i \\ a_k & b_k & c_k \\ a_l & b_l & c_l \end{vmatrix}, \tag{113}$$

where a_i, b_i, c_i are three independent vectors, the dual pseudo-vector is

$$V_i = \frac{1}{i3!}\,\delta_{iklm}\,V_{klm} = \frac{1}{i}\,\delta_{iklm}\,a_k b_l c_m \tag{114}$$

or

$$V_i = \frac{1}{i}\,V_{klm}, \tag{115}$$

where $(iklm)$ is an even permutation of (1234).

V_{ikl} and V_i represent the three-dimensional parallelepiped defined by the vectors a_i, b_i, c_i. V_i is orthogonal to this space, for we get from (114)

$$V_i a_i = V_i b_i = V_i c_i = 0. \tag{116}$$

The volume V of the parallelepiped is given by the length of the pseudo-vector V_i

$$V^2 = -V_i V_i = \frac{1}{3!}\,V_{klm}\,V_{klm}. \tag{117}$$

Finally, the dual pseudo-tensor of an antisymmetrical tensor of rank 4 is a pseudo-invariant. If the tensor is of the form

$$\Sigma_{iklm} = \begin{vmatrix} a_i & b_i & c_i & d_i \\ a_k & b_k & c_k & d_k \\ a_l & b_l & c_l & d_l \\ a_m & b_m & c_m & d_m \end{vmatrix}, \tag{118}$$

the dual pseudo-invariant is

$$\Sigma = \frac{1}{i4!}\delta_{iklm}\Sigma_{iklm} = \frac{1}{i}\delta_{iklm}a_i b_k c_l d_m = \frac{1}{i}\begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}. \tag{119}$$

Σ represents the volume of the parallelepiped defined by the vectors a_i, b_i, c_i, d_i. We have

$$\Sigma^2 = -\frac{1}{4!}\Sigma_{iklm}\Sigma_{iklm}. \tag{120}$$

If a_i, b_i, c_i, d_i are infinitesimal vectors lying in the directions of the coordinate axes and of lengths dx_1, dx_2, dx_3, dx_4, respectively, the corresponding four-dimensional volume element is

$$d\Sigma = \frac{1}{i}dx_1 dx_2 dx_3 dx_4, \tag{121}$$

which is thus a pseudo-invariant.

45. Infinitesimal Lorentz transformations. Lorentz transformations without rotation

An infinitesimal homogeneous linear transformation $(x_i) \to (x_i')$ has the form

$$x_i' = x_i + \epsilon_{ik}x_k = (\delta_{ik}+\epsilon_{ik})x_k, \tag{122}$$

where the ϵ_{ik} are infinitesimal quantities. For a Lorentz transformation we get, using (122) in (10) and neglecting terms of second order in ϵ_{ik},

$$x_i\bar{x}_i = (x_i+\epsilon_{ik}x_k)(\bar{x}_i+\epsilon_{ik}\bar{x}_k) = x_i\bar{x}_i+\bar{x}_i\epsilon_{ik}x_k+x_i\epsilon_{ik}\bar{x}_k.$$

Since this equation must hold for all values of x_i and $\bar{x}_i$ we must have

$$\epsilon_{ik} = -\epsilon_{ki}. \tag{123}$$

This condition is equivalent to the orthogonality relations (11), (14) in the case of an infinitesimal transformation.

Now consider an arbitrary Lorentz transformation (3) connecting the space-time coordinates of two systems S and S'. Let $\mathbf{v} = (v_x, v_y, v_z)$ be the velocity of S' relative to S. The corresponding four-velocity V_i is then by (39)

$$V_i = (\gamma \mathbf{v}, \gamma ic) \quad (\gamma = (1-v^2/c^2)^{-\frac{1}{2}}).$$

The components of this four-vector relative to S' are

$$V'_\iota = (0, 0, 0, ic),$$

since V_ι is the four-velocity of a point at rest in S', which means that the vector V_i lies in the direction of the x'_4-axis. From the transformation equations of a four-vector we then get

$$V_k = V'_\iota \alpha_{\iota k} = ic\alpha_{4k}. \tag{124}$$

Similarly, if $e_\iota^{(1)}$, $e_\iota^{(2)}$, $e_\iota^{(3)}$ are unit vectors in the directions of the x'_1-, x'_2-, x'_3-axes, respectively, we have

$$e_k^{(\iota)'} = \delta_{k\iota} \quad (\iota = 1, 2, 3)$$

and

$$\left.\begin{aligned} e_k^{(\iota)} &= e_l^{(\iota)'}\alpha_{lk} = \delta_{l\iota}\alpha_{lk} = \alpha_{\iota k} \\ e_l^{(\iota)}e_l^{(\kappa)} &= e_l^{(\iota)'}e_l^{(\kappa)'} = \delta_{l\iota}\delta_{lk} = \delta_{\iota\kappa} \\ e_l^{(\iota)}V_l &= e_l^{(\iota)'}V'_l = 0 \end{aligned}\right\}. \tag{124'}$$

In the case of a Lorentz transformation without rotation the coefficients $\alpha_{\iota k}$ follow from (II. 27):

$$\alpha_{\iota k} = \begin{bmatrix} 1+\frac{v_x^2}{v^2}(\gamma-1) & \frac{v_x v_y}{v^2}(\gamma-1) & \frac{v_x v_z}{v^2}(\gamma-1) & \frac{iV_1}{c} \\ \frac{v_y v_x}{v^2}(\gamma-1) & 1+\frac{v_y^2}{v^2}(\gamma-1) & \frac{v_y v_z}{v^2}(\gamma-1) & \frac{iV_2}{c} \\ \frac{v_z v_x}{v^2}(\gamma-1) & \frac{v_z v_y}{v^2}(\gamma-1) & 1+\frac{v_z^2}{v^2}(\gamma-1) & \frac{iV_3}{c} \\ \frac{-iV_1}{c} & \frac{-iV_2}{c} & \frac{-iV_3}{c} & \frac{-iV_4}{c} \end{bmatrix}. \tag{125}$$

In $(3+1)$-space it represents a rotation in the two-dimensional plane defined by the time axes of S' and S.

If v_x, v_y, v_z are infinitesimal quantities, we have, neglecting second-order terms in these quantities,

$$V_1 = v_x, \quad V_2 = v_y, \quad V_3 = v_z, \quad V_4 = ic,$$

and (125) reduces to

$$\alpha_{\iota k} = \delta_{\iota k}+\epsilon_{ik} \quad \text{with} \quad \epsilon_{\iota\kappa} = 0, \quad \epsilon_{\iota 4} = -\epsilon_{4\iota} = \frac{iV_\iota}{c}, \quad \epsilon_{44} = 0. \tag{126}$$

46. Successive Lorentz transformations

Let

$$x'_\iota = \alpha_{\iota k} x_k, \quad S \to S', \qquad x''_i = \alpha'_{\iota k} x'_k, \quad S' \to S'' \tag{127}$$

be two successive Lorentz transformations. The resultant transformation

$$x''_\iota = (\alpha'_{\iota l}\alpha_{lk})x_k \tag{128}$$

is, of course, again a Lorentz transformation, i.e. the coefficients

$$\alpha''_{ik} = \alpha'_{il}\alpha_{lk}$$

satisfy the same orthogonality relations (11) and (14) as α_{ik} and α'_{ik}. However, (128) will not in general represent a Lorentz transformation without rotation even if this is the case for the two transformations (127), i.e. α''_{ik} will not take the form (125) even if α_{ik} and α'_{ik} are of this form. This will be the case only if the three time axes in S, S', and S'' are lying in the same plane.

In the special case where the transformation from S' to S'' is an infinitesimal Lorentz transformation without rotation, we have, according to (126),

$$\alpha'_{ik} = \delta_{ik}+\epsilon'_{ik}, \qquad \epsilon'_{\iota\kappa} = 0, \qquad \epsilon'_{\iota 4} = -\epsilon'_{4\iota} = \frac{iV'_\iota}{c}, \qquad \epsilon'_{44} = 0, \tag{129}$$

V'_ι being the four-velocity of S'' relative to S'. Hence,

$$\alpha''_{ik} = (\delta_{il}+\epsilon'_{il})\alpha_{lk} = \alpha_{ik}+\epsilon'_{il}\alpha_{lk}. \tag{130}$$

Let

$$x_i = f_\iota(\tau)$$

represent the time track of a particle in arbitrary motion in S, τ being the proper time of the particle. We shall now try to determine the successive rest systems of the particle such that two consecutive rest systems at any time have the same orientation of the spatial axes. Let S' and S'' in (127) be momentary rest systems of the particle at the times τ and $\tau+d\tau$, respectively. The four-velocity of S' relative to S is then

$$U_\iota(\tau) = \frac{dx_i}{d\tau} = \dot{f}_\iota(\tau). \tag{131}$$

Similarly, the four-velocity of S'' relative to S is

$$V_i = U_\iota(\tau)+dU_\iota(\tau) = U_\iota(\tau)+\dot{U}_i\,d\tau = \dot{f}_i(\tau)+\ddot{f}_i(\tau)\,d\tau. \tag{132}$$

The components of these two four-velocities in S' are

$$U'_i(\tau) = \alpha_{ik}U_k = (0,0,0,ic), \tag{133}$$

$$V'_i = U'_\iota+dU'_i = \alpha_{ik}(U_k+dU_k) = U'_i+\alpha_{ik}\,dU_k. \tag{134}$$

Since the transformation from S' to S'' was supposed to be an infinitesimal Lorentz transformation without rotation, the coefficients α''_{ik} in the transformation from S to S'' are obtained from (130) and (129), V'_ι in (129) being given by (134). Now, obviously, we have

$$\epsilon'_{ik} = \frac{V'_i U'_k - V'_k U'_i}{c^2} = \frac{U'_k dU'_i - U'_i dU'_k}{c^2}, \tag{135}$$

for, on account of (133), this is seen to be identical with the ϵ'_{ik} defined in (129).

Thus, from (130), (135), and (134)

$$\alpha''_{ik}-\alpha_{ik} = \frac{U'_l\,dU'_i - U'_i\,dU'_l}{c^2}\alpha_{lk} = \frac{1}{c^2}\alpha_{il}(U_k\,dU_l - U_l\,dU_k).$$

The coefficients α_{ik} can now be regarded as functions $\alpha_{ik}(\tau)$ of τ, α''_{ik} being then equal to $\alpha_{ik}(\tau+d\tau)$. Hence we get the following differential equations for the functions $\alpha_{ik}(\tau)$:

$$\frac{d\alpha_{ik}(\tau)}{d\tau} = \alpha_{il}\,\eta_{lk} \tag{136}$$

with

$$\eta_{ik} = \frac{\dot{U}_i\,U_k - \dot{U}_k\,U_i}{c^2}. \tag{137}$$

For later use we note that the coefficients α_{ik} and η_{ik} satisfy the relations

$$\left.\begin{aligned} \alpha_{\kappa i}\,\alpha_{\lambda i} &= \delta_{\kappa\lambda}, & \alpha_{\kappa i}\,U_i &= 0 \\ \alpha_{\lambda l}\,\eta_{li}\,\alpha_{\kappa i} &= 0, & \alpha_{\kappa l}\,\eta_{li}\,U_i &= -\alpha_{\kappa l}\,\dot{U}_l \\ \alpha_{\kappa l}\,\eta_{li}\,\alpha_{\lambda m}\,\eta_{mi} &= -\frac{(\alpha_{\kappa l}\,\dot{U}_l)(\alpha_{\lambda m}\,\dot{U}_m)}{c^2} \end{aligned}\right\} \tag{138}$$

on account of (14), (41), (41′) and (133).

The equations (136) determine the transformation from the fixed system S to the momentary rest system $S' = S'(\tau)$ with coordinates x'_i, thus we have

$$x_i = x'_k\,\alpha_{ki}(\tau), \tag{139}$$

or, if we make a continuous displacement of the origin in S' such that the particle is always lying at the origin of the rest systems $S'(\tau)$,

$$x_i = f_i(\tau) + x'_k\,\alpha_{ki}(\tau). \tag{140}$$

Let us now attach a space vector $\mathbf{e}'(\tau)$ of unit length to the particle considered in such a way that the components $\mathbf{e}'$ with respect to the spatial axes of $S'(\tau)$ have the same values at all times. $\mathbf{e}'(\tau)$ may, for instance, have the direction of one of the space axes in S'. We thus have at any time a displacement of $\mathbf{e}'$ without change of orientation. In (3+1)-space this vector is represented by a space-like four-vector with components given by

$$e'_i = (\mathbf{e}', 0) \tag{141}$$

in S'. Its components in S are given by

$$e_i(\tau) = e'_k\,\alpha_{ki}(\tau) = e'_\kappa\,\alpha_{\kappa i}(\tau). \tag{142}$$

e_i is orthogonal to U_i, since

$$e_i\,U_i = e'_i\,U'_i = 0 \tag{143}$$

on account of (133) and (141). From (142), (136), (137), and (143) we now have

$$\frac{de_\iota(\tau)}{d\tau} = e'_k \frac{d\alpha_{ki}(\tau)}{d\tau} = e'_k \alpha_{kl} \frac{\dot{U}_l U_\iota - \dot{U}_i U_l}{c^2}$$

$$= \frac{(e_l \dot{U}_l)U_\iota - \dot{U}_\iota(e_l U_l)}{c^?},$$

i.e.

$$\frac{de_\iota(\tau)}{d\tau} = \frac{(e_l \dot{U}_l)U_\iota}{c^2}. \tag{144}$$

If $S'(\tau)$ coincides with S for $\tau = 0$ we have

$$e_i = e'_\iota \quad \text{for} \quad \tau = 0$$

and the velocity of the particle is zero at that moment. At a later time we have in general $e_\iota \neq e'_\iota$, and even if the velocity of the particle becomes zero again, we shall in general have

$$e_\iota = (\mathbf{e}, 0) \neq e'_\iota = (\mathbf{e}', 0)$$

at that time. This means that the components of the unit vector considered are different in S and S', in accordance with the fact that the vector has performed a *Thomas precession* relative to S (see § 22).

47. Successive rest systems of a particle in arbitrary rectilinear motion and in constant circular motion

Let the motion of the particle be in the direction of the x_1-axis, then we have $f_2 = f_3 = 0$, i.e.

$$U_i = (\dot{f}_1, 0, 0, \dot{f}_4), \qquad \dot{U}_\iota = (\ddot{f}_1, 0, 0, \ddot{f}_4). \tag{145}$$

Further, we shall assume that the particle has zero velocity in S at the time $\tau = 0$ and that $S'(\tau)$ coincides with S at that time. On account of the equation $U_\iota U_\iota = U_1^2 + U_4^2 = -c^2$ we may therefore write U_ι in the form

$$U_\iota = (c \sinh\theta(\tau), \quad 0, \quad 0, \quad ic\cosh\theta(\tau)), \tag{146}$$

where $\theta(\tau)$ is an arbitrary function of τ which is zero for $\tau = 0$. Hence,

$$f_\iota(\tau) = \left(c\int_0^\tau \sinh\theta\, d\tau, \quad 0, \quad 0, \quad ic\int_0^\tau \cosh\theta\, d\tau\right). \tag{147}$$

Further, if $e_k^{(\iota)}(\tau)$ represent unit vectors in the directions of the spatial axes of the successive rest systems $S'(\tau)$ we have

$$e_k^{(\iota)}(0) = \delta_{k\iota} \quad \text{for} \quad \tau = 0. \tag{148}$$

Each of these vectors satisfies the equations (144);

therefore

$$\frac{de_k^{(\iota)}}{d\tau} = (e_l^{(\iota)}\dot{U}_l)U_k/c^2. \tag{149}$$

On account of (145) we see at once that

$$e_k^{(2)} = \delta_{k2}, \quad e_k^{(3)} = \delta_{k3}, \quad e_2^{(1)} = e_3^{(1)} = 0 \tag{150}$$

are solutions of (149). To find the components $e_1^{(1)}$ and $e_4^{(1)}$ we use the circumstance that

$$e_k^{(1)}U_k = 0, \quad e_k^{(1)}e_k^{(1)} = 1 \tag{151}$$

are integrals of the equations (149), as is seen at once by multiplying (149) by U_k and $e_k^{(1)}$ respectively. Hence

$$\left.\begin{aligned} e_1^{(1)}U_1+e_4^{(1)}U_4 &= 0, \qquad e_4^{(1)} = -e_1^{(1)}\frac{U_1}{U_4}, \\ e_1^{(1)2}+e_4^{(1)2} &= 1, \\ e_1^{(1)2}\left(1+\frac{U_1^2}{U_4^2}\right) &= -e_1^{(1)2}\frac{c^2}{U_4^2} = 1 \end{aligned}\right\}, \tag{152}$$

i.e.

$$\left.\begin{aligned} e_1^{(1)} &= \frac{U_4}{ic}, \\ e_4^{(1)} &= -\frac{U_4}{ic}\frac{U_1}{U_4} = \frac{i}{c}U_1 \end{aligned}\right\}. \tag{153}$$

From (124) and (124′) we therefore simply get

$$\alpha_{ik} = \begin{pmatrix} U_4/ic & 0 & 0 & iU_1/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ U_1/ic & 0 & 0 & U_4/ic \end{pmatrix}, \tag{154}$$

which correspond to special Lorentz transformations (see (16)).

By means of (154),. (146), and (147) the transformation equations (140) may be written

$$\left.\begin{aligned} x_1 &= c\int\limits_0^\tau \sinh\theta\, d\tau + x_1'\cosh\theta(\tau) + x_4'\frac{\sinh\theta(\tau)}{i} \\ x_2 &= x_2', \qquad x_3 = x_3' \\ x_4 &= ic\int\limits_0^\tau \cosh\theta\, d\tau + x_1'\, i\sinh\theta(\tau) + x_4'\cosh\theta(\tau) \end{aligned}\right\}. \tag{155}$$

In the special case where the motion of the particle is hyperbolic we get from (III. 46)

$$d\tau = \sqrt{\left\{1-\frac{u^2}{c^2}\right\}}\, dt = \frac{dt}{\sqrt{\{1+(gt/c)^2\}}},$$

i.e.

$$\tau = \frac{c}{g}\sinh^{-1}\frac{gt}{c}, \qquad t = \frac{c}{g}\sinh\frac{g\tau}{c}. \tag{156}$$

Thus, from (III. 47) and (156)

$$\left.\begin{aligned} f_1(\tau) &= \frac{c^2}{g}\left(\sqrt{\left\{1+\frac{g^2t^2}{c^2}\right\}}-1\right) = \frac{c^2}{g}\left(\cosh\frac{g\tau}{c}-1\right) \\ f_2 &= 0, \qquad f_3 = 0 \\ f_4(\tau) &= i\frac{c^2}{g}\sinh\frac{g\tau}{c} \end{aligned}\right\}, \tag{157}$$

$$\left.\begin{aligned} U_i = f_i &= \left(c\sinh\frac{g\tau}{c}, 0, 0, ic\cosh\frac{g\tau}{c}\right) \\ \dot{U}_i &= \left(g\cosh\frac{g\tau}{c}, 0, 0, ig\sinh\frac{g\tau}{c}\right) \end{aligned}\right\}. \tag{158}$$

A comparison with (146) shows that $\theta(\tau)$ in this particular case is

$$\theta(\tau) = \frac{g\tau}{c}, \tag{159}$$

and the transformation (155) reduces to

$$\left.\begin{aligned} x_1 &= \frac{c^2}{g}\left(\cosh\frac{g\tau}{c}-1\right)+x_1'\cosh\frac{g\tau}{c}+x_4'\frac{\sinh g\tau/c}{i} \\ x_2 &= x_2', \qquad x_3 = x_3' \\ x_4 &= i\frac{c^2}{g}\sinh\frac{g\tau}{c}+x_1' i\sinh\frac{g\tau}{c}+x_4'\cosh\frac{g\tau}{c} \end{aligned}\right\}. \tag{160}$$

By means of the transformation coefficients (154) we get for the components $\dot{U}_i'$ of the four acceleration vectors of the particle in the successive rest system $S'(\tau)$

$$\dot{U}_i' = \alpha_{ik}\dot{U}_k = \left(\frac{U_4\dot{U}_1}{ic}+\frac{i}{c}U_1\dot{U}_4, 0, 0, \frac{U_1\dot{U}_1+U_4\dot{U}_4}{ic}\right),$$

or, on account of (158),

$$\dot{U}_i' = (g, 0, 0, 0), \tag{161}$$

which shows that the accelerations of the particle in the successive rest systems are constantly equal to g (see equation (42′)).

We shall now briefly discuss the solution of the equations (136) in the case where the particle is moving in the $(x_1 x_2)$-plane with constant angular velocity ω in a circle of radius a. In this case we have

$$f_i = (a\cos\omega\gamma\tau,\ a\sin\omega\gamma\tau,\ 0, ic\gamma\tau), \tag{162}$$

with $\gamma = (1-u^2/c^2)^{-\frac{1}{2}}$, $u = a\omega$ being the constant velocity of the particle in the circular motion. From (162) we get

$$\left.\begin{aligned} U_i &= (-a\omega\gamma\sin(\omega\gamma\tau), a\omega\gamma\cos(\omega\gamma\tau), 0, ic\gamma) \\ \dot{U}_i &= (-a\omega^2\gamma^2\cos(\omega\gamma\tau), -a\omega^2\gamma^2\sin(\omega\gamma\tau), 0, 0) \end{aligned}\right\}. \tag{163}$$

By a straightforward calculation it is easily verified that the following scheme of coefficients α_{ik} provides a solution of the equations (136):

$$\alpha_{ik} = \begin{bmatrix} \cos\alpha\cos\beta+\gamma\sin\alpha\sin\beta & \sin\alpha\cos\beta-\gamma\cos\alpha\sin\beta & 0 & -i\frac{u\gamma}{c}\sin\beta \\ \cos\alpha\sin\beta-\gamma\sin\alpha\cos\beta & \sin\alpha\sin\beta+\gamma\cos\alpha\cos\beta & 0 & i\frac{u\gamma}{c}\cos\beta \\ 0 & 0 & 1 & 0 \\ i\frac{u\gamma}{c}\sin\alpha & -i\frac{u\gamma}{c}\cos\alpha & 0 & \gamma \end{bmatrix}, \tag{164}$$

with $\alpha = \omega\gamma\tau$, $\beta = \gamma\alpha = \omega\gamma^2\tau$.

The transformation $S \rightarrow S'(\tau)$ is then obtained from (139) or (140) by introducing the expressions (164) for α_{ik}.

For $\tau = 0$ we get

$$\alpha_{ik}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \gamma & 0 & i\frac{u\gamma}{c} \\ 0 & 0 & 1 & 0 \\ 0 & -i\frac{u}{c}\gamma & 0 & \gamma \end{bmatrix}. \tag{165}$$

Thus, denoting the space-time coordinates of the system $S'(0)$ by x_i^0, we get from (139)

$$\left.\begin{aligned} x_1 &= x_1^0, \quad x_2 = \gamma x_2^0 - i\frac{u\gamma}{c}x_4^0 \\ x_3 &= x_3^0, \quad x_4 = i\frac{u\gamma}{c}x_2^0+\gamma x_4^0 \end{aligned}\right\}, \tag{166}$$

which represents a special Lorentz transformation from S to a system $S'(0)$ moving with the velocity u in the direction of the x_2-axis.

At the later time $\tau = \tau_1 = 2\pi/\omega\gamma$ we get from (164)

$$\alpha_{ik}(\tau_1) = \begin{bmatrix} \cos\beta_1 & -\gamma\sin\beta_1 & 0 & -i\frac{u\gamma}{c}\sin\beta_1 \\ \sin\beta_1 & \gamma\cos\beta_1 & 0 & i\frac{u\gamma}{c}\cos\beta_1 \\ 0 & 0 & 1 & 0 \\ 0 & -i\frac{u\gamma}{c} & 0 & \gamma \end{bmatrix}, \tag{167}$$

where $\beta_1 = 2\pi(\gamma-1)$. The coefficients $\alpha_{ik}(\tau_1)$ may, however, be written as

$$\alpha_{ik}(\tau_1) = \beta_{il}\,\alpha_{lk}(0),$$

where
$$\beta_{ik} = \begin{pmatrix} \cos\beta_1 & -\sin\beta_1 & 0 & 0 \\ \sin\beta_1 & \cos\beta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{168}$$

If the space-time coordinates of the system $S'(\tau_1)$ are denoted by x_i^1 we have
$$x_i^1 = \alpha_{ik}(\tau_1)x_k = \beta_{il}\alpha_{lk}(0)x_k = \beta_{il}x_l^0$$
or
$$\begin{aligned} x_1^1 &= \cos\beta_1\, x_1^0 - \sin\beta_1\, x_2^0, \\ x_2^1 &= \sin\beta_1\, x_1^0 + \cos\beta_1\, x_2^0, \\ x_3^1 &= x_3^0, \qquad x_4^1 = x_4^0. \end{aligned} \tag{169}$$

Thus the system $S'(\tau_1)$ does not coincide with the system $S'(0)$, but has to be rotated through an angle β_1 in the $(x_1 x_2)$-plane in the direction of motion of the particle in order to give the spatial axes of $S'(\tau_1)$ the same orientation as the axes of $S'(0)$ or as the axes of S. In other words, the axes of S have to be rotated through an angle $-\beta_1 = -2\pi(\gamma-1)$ in the $(x_1 x_2)$-plane in order to give them the same orientation as the axis of the system $S(\tau_1)$. This is due to the Thomas effect. Integrating the formula (II. 65) for the velocity of the Thomas precession over a whole period T, we get
$$\int\limits_0^T \boldsymbol{\omega}\, dt = -\int\limits_0^T (\gamma-1)\frac{\mathbf{v}\times\dot{\mathbf{v}}}{v^2}\, dt$$

and, since $\mathbf{v}\times\dot{\mathbf{v}}$ in our case is a constant vector perpendicular to $\mathbf{v}$ and $\dot{\mathbf{v}}$ and $\dot{v}T = v\omega T = 2\pi v$, the total angle of precession is just
$$\theta = -2\pi(\gamma-1) = -\beta_1. \tag{170}$$

48. Tensor and pseudo-tensor fields. Tensor analysis

As in ordinary space we speak of a tensor field of rank n in $(3+1)$-space if to any point in this space is connected a tensor of rank n. In particular, we have a tensor field of rank zero, a so-called scalar field, if an invariant number is connected with every event point. This means that we have a certain function of the coordinates $\phi(x) = \phi(x_i) = \phi(x_1, x_2, x_3, x_4)$ in every system of coordinates S, such that
$$\phi'(x') = \phi(x), \tag{171}$$

if $\phi'(x')$ is the function corresponding to the system of coordinates S' and the numbers (x_i') and (x_i) are coordinates of the *same* event point in the two systems S' and S, respectively. In general, ϕ' will be a different function of the variables (x_i') than ϕ is of (x_i). Thus a scalar function $\phi(x_i)$ is generally *not* a form-invariant function of the coordinates. This will in fact only be the case if ϕ is a function merely of the quantity (4) which is also invariant in form.

Analogously, we have a tensor field of rank 1 when a four-vector is connected with every event point. The components $a_i(x)$ and $a'_i(x')$ of the four-vector in two arbitrary systems of coordinates S and S' will then be functions of the coordinates of the event points, and

$$a'_i(x') = \alpha_{ik}\, a_k(x) \tag{172}$$

when the connexion between the variables $(x') = (x'_i)$ and $(x) = (x_i)$ is given by (3) and (13).

For tensor fields of higher rank equations exactly analogous to (171) and (172) are valid.

Now from an arbitrary scalar field $\phi(x)$ we can by means of a covariant operation form a vector field with the components $\partial\phi/\partial x_i$ in the arbitrary system of coordinates S; for, from (171), we obtain by differentiation

$$\frac{\partial\phi'}{\partial x'_i} = \frac{\partial\phi}{\partial x_k}\frac{\partial x_k}{\partial x'_i} = \alpha_{ik}\frac{\partial\phi}{\partial x_k} \tag{173}$$

where we have used
$$\frac{\partial x_k}{\partial x'_i} = \alpha_{ik}, \tag{174}$$
following from (13).

The four-vector $\partial\phi/\partial x_i$ is called the gradient of ϕ and is written

$$\operatorname{grad}_i \phi = \frac{\partial\phi}{\partial x_i}. \tag{175}$$

It is analogous to the ordinary gradient vector in three dimensions.

Similarly, from a vector field $a_i(x)$ we can form a tensor field of rank 2 with the components $\partial a_i/\partial x_k$ in an arbitrary system of coordinates, for by differentiation of (172) we get

$$\frac{\partial a'_i}{\partial x'_k} = \alpha_{il}\frac{\partial a_l}{\partial x'_k} = \alpha_{il}\frac{\partial x_m}{\partial x'_k}\frac{\partial a_l}{\partial x_m} = \alpha_{il}\alpha_{km}\frac{\partial a_l}{\partial x_m}. \tag{176}$$

The antisymmetrical combination $\partial a_k/\partial x_i - \partial a_i/\partial x_k$ is also a tensor field of rank 2 which is called the curl of the vector field $a_i(x)$. It is denoted by

$$\operatorname{curl}_{ik} a \equiv \operatorname{curl}_{ik}\{a_i\} \equiv \frac{\partial a_k}{\partial x_i} - \frac{\partial a_i}{\partial x_k}. \tag{177}$$

(In the expression $\operatorname{curl}_{ik}\{a_i\}$ we shall of course not sum over i.) The curl is an antisymmetrical tensor. Therefore, in three-dimensional space it corresponds to an axial vector, viz. to $\operatorname{curl} \mathbf{a}$.

By contraction of the tensor field $\partial a_\iota/\partial x_k$ we obtain a tensor of rank zero. From a vector field $a_\iota(x)$ we can thus construct a scalar field

$$\frac{\partial a_i}{\partial x_i} = \frac{\partial a'_\iota}{\partial x'_\iota}, \tag{178}$$

which is called the divergence of the vector field a_i. It is denoted by

$$\operatorname{div} a \equiv \operatorname{div}\{a_i\} = \frac{\partial a_\iota}{\partial x_\iota} \tag{179}$$

and is analogous to the ordinary three-dimensional divergence div **a**.

If a_ι is the gradient of a scalar ψ, i.e.

$$a_\iota = \frac{\partial \psi}{\partial x_\iota}, \tag{180}$$

the divergence of a_ι becomes

$$\frac{\partial a_\iota}{\partial x_i} = \frac{\partial^2 \psi}{\partial x_\iota\, \partial x_\iota} = \Box \psi, \tag{181}$$

where we have put

$$\Box = \frac{\partial^2}{\partial x_\iota\, \partial x_\iota} = \frac{\partial^2}{\partial x_\iota\, \partial x_\iota} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} = \Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{182}$$

The operator (182), d'Alembert's operator, is thus a covariant operator; it is the four-dimensional generalization of Laplace's operator

$$\Delta = \frac{\partial^2}{\partial x_\iota\, \partial x_\iota}.$$

In the same way, by means of differentiations we can always form a tensor field of rank $n+1$ from a tensor field of rank n, and by subsequent contraction we then get a tensor field of rank $n-1$. As in the special cases just considered, this is a consequence of the transformation equations for tensors together with the equation (173) and the orthogonality relations (11) and (14). From a tensor field t_{ik} of rank 2 we can thus construct a tensor field $\partial t_{ik}/\partial x_l$ of rank 3, and by contraction we obtain a tensor field of rank 1, i.e. a vector field

$$\operatorname{div}_i t \equiv \operatorname{div}_\iota\{t_{\iota k}\} = \frac{\partial t_{\iota k}}{\partial x_k}, \tag{183}$$

which is called the divergence of the tensor field t_{ik}.

If the tensor field F_{ik} is antisymmetrical, we can obviously form a completely antisymmetrical tensor field of rank 3

$$\operatorname{curl}_{\iota kl} F \equiv \operatorname{curl}_{\iota kl}\{F_{ik}\} = \frac{\partial F_{\iota k}}{\partial x_l} + \frac{\partial F_{kl}}{\partial x_i} + \frac{\partial F_{li}}{\partial x_k}, \tag{184}$$

which is called the curl of the tensor F_{ik}. If we interchange any two of the three independent indices in (184) the expression changes sign. Thus, $\text{curl}_{ikl}F$ is zero if two of the indices are equal and the tensor curl_{ikl} has only $\frac{4!}{3!\,1!} = 4$ independent components. If F_{ik} is equal to the curl of a vector field, i.e. if

$$F_{ik} = \text{curl}_{ik}\, a = \frac{\partial a_k}{\partial x_i} - \frac{\partial a_i}{\partial x_k}, \tag{185}$$

the curl of F_{ik} is identically zero, i.e.

$$\text{curl}_{ikl} F \equiv 0. \tag{186}$$

In the same way we speak of a pseudo-tensor field when to each point in space is connected a pseudo-tensor. These are also frequently called tensor densities, since $t_{ik..}\, dx_1 dx_2 dx_3 dx_4$, according to (121), is a tensor when $t_{ik.}$ is a pseudo-tensor. Thus a pseudo-scalar is a tensor density of rank zero, a pseudo-vector a tensor density of rank 1, etc. To the various antisymmetrical tensor fields can be attributed dual tensor densities which have the same reciprocal connexion as have the tensors and their dual pseudo-tensors defined in § 44. The vector density dual to $\text{curl}_{ikl}F$ is equal to $\text{div}_i\, F^*$, where F^*_{ik} is the tensor density dual to F_{ik}, for, according to (108), we have

$$\frac{\partial F^*_{ik}}{\partial x_k} = \frac{1}{2i}\delta_{iklm}\frac{\partial F_{lm}}{\partial x_k} = \frac{1}{i3!}\delta_{iklm}\left(\frac{\partial F_{lm}}{\partial x_k} + \frac{\partial F_{mk}}{\partial x_l} + \frac{\partial F_{kl}}{\partial x_m}\right), \tag{187}$$

which just means that $\text{div}_i\, F^*$ and $\text{curl}_{ikl}\, F$ are dual to each other

49. Gauss's theorem in four-dimensional space

If $\mathbf{a} = \mathbf{a}(\mathbf{x})$ is a three-vector field and V a domain in 3-space bounded by a closed surface σ, Gauss's theorem in ordinary space is expressed by the equation

$$\int_V \text{div}\,\mathbf{a}\, dV = \int_\sigma a_n\, d\sigma, \tag{188}$$

where a_n is the component of $\mathbf{a}$ in the direction of the outward normal n to the surface element $d\sigma$. (An elementary proof of this theorem is given in Appendix 1). Gauss's theorem thus permits of transforming the volume integral on the left-hand side of (188) into an integral over the two-dimensional boundary σ of the volume V. If $\mathbf{n}$ is a unit vector in the direction of the outward normal, (188) may also be written

$$\int_V \frac{\partial a_i}{\partial x_i} dV = \int_\sigma a_i n_i\, d\sigma, \tag{189}$$

where dV is given by (107).

If we choose the surface element in the form of a parallelogram formed by infinitesimal vectors dx_ι and δx_ι lying in the surface σ, the surface element may be represented by the tensor $d\sigma_{\iota\kappa} = dx_\iota \delta x_\kappa - dx_\kappa \delta x_\iota$ obtained from (99) by substituting dx_ι and δx_ι for a_ι and b_ι, respectively. Alternatively the infinitesimal parallelogram may also be represented by the corresponding axial vector $d\sigma_\iota$ defined by (100). Since this vector is perpendicular to the surface element, (189) can also be written in the form

$$\int_V \frac{\partial a_\iota}{\partial x_\iota}\, dV = \int_\sigma a_\iota\, d\sigma_\iota = \int a_\iota\, \delta_{\iota\kappa\lambda}\, dx_\kappa \delta x_\lambda \tag{190}$$

provided that the sequence of the vectors dx_ι and δx_ι is chosen so that the axial vector $d\sigma_\iota$ lies in the direction of the normal pointing away from the domain V.

In this form Gauss's theorem may be immediately generalized to four dimensions.† If $a_i(x)$ is a four-vector field and Σ a domain in (3+1)-space bounded by the closed three-dimensional surface V, Gauss's generalized theorem takes the form

$$\int_\Sigma \frac{\partial a_i}{\partial x_i}\, d\Sigma = \int_V a_\iota\, dV_\iota = \frac{1}{i} \int a_\iota\, \delta_{iklm}\, dx_k \delta x_l\, \Delta x_m, \tag{191}$$

where $d\Sigma$ is given by (121) and (dx_ι), (δx_i), and (Δx_i) are three four-vectors lying in the boundary space V. The pseudo-four-vector dV_i is given by (114) with (a_ι), (b_ι), (c_i) equal to (dx_ι), (δx_ι), (Δx_ι), respectively.

If a part of the boundary V is a hyperplane Ω defined by $x_4 =$ constant, the vectors (dx_ι), (δx_ι), (Δx_ι) are orthogonal to the time axis and we can choose

$$dx_\iota = (dx_1, 0, 0, 0),$$
$$\delta x_i = (0, dx_2, 0, 0),$$
$$\Delta x_i = (0, 0, dx_3, 0).$$

The pseudo-vector dV_i on the hyperplane Ω will then have the components

$$dV_i = (0, 0, 0, \pm i\, dx_1 dx_2 dx_3), \tag{191'}$$

where the plus or the minus sign should be taken according as the normal to Ω, pointing away from the region Σ, lies in the direction of the negative or the positive time axis.

If $t_{\iota\kappa}$ is a space tensor, we have by analogy with (189) and (190)

$$\int_V \frac{\partial t_{\iota\kappa}}{\partial x_\kappa}\, dV = \int_\sigma t_{\iota\kappa} n_\kappa\, d\sigma = \int_\sigma t_{\iota\kappa}\, d\sigma_\kappa = \int_\sigma t_{\iota\kappa}\, \delta_{\kappa\lambda\mu}\, dx_\lambda \delta x_\mu, \tag{192}$$

† A. Sommerfeld, *Ann. d. Phys.* **32**, 749 (1910); **33**, 649 (1910).

and the analogous equation in (3+1)-space is

$$\int\limits_{\Sigma} \frac{\partial t_{ik}}{\partial x_k}\, d\Sigma = \int\limits_{V} t_{ik}\, dV_k = \int t_{ik}\, \delta_{klmn}\, dx_l\, \delta x_m\, \Delta x_n. \tag{193}$$

50. The fundamental equations of mechanics for incoherent matter

As a first application of the mathematical methods developed in the preceding sections we shall now consider the motion of continuously distributed matter under the influence of given external forces. In order to be able to apply to such a system the fundamental equations of motion of material particles, deduced in Chapter III, we shall regard a continuous mass distribution as a limiting case of a distribution of a very large number of material particles. If the particles are so small and the number of particles per unit volume is so large that our macroscopic measuring instruments cannot distinguish between the single particles, the mass distribution in an arbitrary system of inertia S can be described by a mass density $\mu(\mathbf{x}, t)$ which, for practical purposes, may be regarded as a continuous function of the space and time variables. $\mu(\mathbf{x}, t)$ is defined so that $\mu\, \delta V$ is equal to the total mass inside the volume element δV at the place $\mathbf{x}$ and at the time t. The motion of the matter at any place and at any time is described by a velocity vector $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$, which is a function of $\mathbf{x}$ and t Then the mass current density is equal to $\mu\mathbf{u}$. In Newtonian mechanics the mass is a quantity which is conserved. According to formula (III. 22) this is not so in the theory of relativity, when the matter is acted upon by forces, since the velocity of a given material particle, and thus its relativistic mass, in this case changes with time. On the other hand, the proper mass of the matter is conserved in many cases, viz. when the four-force at any point is orthogonal to the four-velocity of the matter at the place considered, i.e. when the equation (57) holds everywhere and at any time.

The density of proper mass μ_0, i.e. the proper mass per unit volume, in the arbitrary system of inertia S is, according to (III. 22), connected with μ by

$$\mu_0 = \mu\sqrt{(1-u^2/c^2)}. \tag{194}$$

μ_0, like μ, is a given function of $\mathbf{x}$ and t.

Now, considering a certain point inside the matter at a given time, we can introduce the system of inertia S^0 relative to which the matter at this point is momentarily at rest. In S^0 (194) reduces to

$$\mu_0^0 = \mu^0, \tag{195}$$

where all quantities referring to S^0 are provided with the superscript 0.

In the rest system the proper mass density is thus identical with the relativistic mass density. In contradistinction to μ_0, $\mu_0^0 = \mu^0$ is thus an invariant. Further, it is also easily seen that $\mu_0\surd(1-u^2/c^2)$ is an invariant. For, if we consider a small piece of matter which in S has a volume δV, in the rest system S^0 it has a volume δV^0 which, according to (II. 34), is connected with δV by the equation

$$\delta V = \delta V^0\surd(1-u^2/c^2). \tag{196}$$

The invariant proper mass of the material particle expressed in the two systems is then given by

$$\mu_0\,\delta V = \mu_0^0\,\delta V^0 = \mu^0\,\delta V^0, \tag{197}$$

which, by means of (196), leads to the equation

$$\mu_0\surd(1-u^2/c^2) = \mu^0 = \text{invariant}, \tag{198}$$

or, according to (194), to
$$\mu\left(1-\frac{u^2}{c^2}\right) = \mu^0. \tag{199}$$

Consequently the quantities on the left-hand side of (198) and (199) must be invariants.

Let now $\phi = \phi(\mathbf{x},t)$ be a given function of the space and time coordinates in S. We must then distinguish between the local time differentiation $\partial\phi/\partial t$ indicating the change of ϕ per unit time at a fixed point in space and the substantial differential coefficient $d\phi/dt$ indicating the change per unit time when we follow the matter in its motion. We obviously have

$$\frac{d\phi}{dt} = \frac{\partial\phi}{\partial t}+(\mathbf{u}\,.\,\mathrm{grad}\,\phi). \tag{200}$$

In the same way we have for a three-dimensional vector field $\mathbf{a} = \mathbf{a}(\mathbf{x},t)$

$$\frac{d\mathbf{a}}{dt} = \frac{\partial\mathbf{a}}{\partial t}+(\mathbf{u}\,.\,\mathrm{grad})\mathbf{a}, \tag{201}$$

since an equation of the type (200) must hold for each component of $\mathbf{a}$.

Let us now consider the matter which at the time t is situated inside a closed surface σ enclosing a domain with the volume V. At the time $t+dt$ the volume of this material will have been increased by an amount

$$dV = dt\int\limits_\sigma u_n\,d\sigma, \tag{202}$$

where the integral on the right-hand side is a surface integral over the surface σ, and where u_n denotes the component of $\mathbf{u}$ in the direction of the outward normal to the surface element $d\sigma$; for every face element $d\sigma$ moves a distance of $\mathbf{u}\,dt$ during the time dt and therefore sweeps through a volume $u_n\,dtd\sigma$ during this time. By means of Gauss's theorem (188) equation (202) can be written

$$\frac{dV}{dt} = \int\limits_V \operatorname{div}\mathbf{u}\,dV, \tag{203}$$

an equation which must hold for every part of the material body. Considering in particular the matter which at the time t lies inside an infinitesimal volume element δV, we get for the volume dilatation per unit time

$$\frac{1}{\delta V}\frac{d\delta V}{dt} = \operatorname{div}\mathbf{u}. \tag{204}$$

Let us begin with the consideration of a system in which the proper mass is conserved. For any material particle with the volume δV at the arbitrary time t we must then have

$$\frac{d}{dt}(\mu_0\,\delta V) = \frac{d\mu_0}{dt}\delta V+\mu_0\frac{d\delta V}{dt} = 0, \tag{205}$$

which, by means of (204), can be written

$$\frac{d\mu_0}{dt}+\mu_0\operatorname{div}\mathbf{u} = 0. \tag{206}$$

This equation which expresses the conservation of the proper mass can also be given another form when we apply equation (200) to the function $\mu_0(\mathbf{x},t)$. We then obtain

$$\frac{\partial\mu_0}{\partial t}+(\mathbf{u}\,.\operatorname{grad}\mu_0)+\mu_0\operatorname{div}\mathbf{u} = 0,$$

or

$$\frac{\partial\mu_0}{\partial t}+\operatorname{div}(\mu_0\,\mathbf{u}) = 0. \tag{207}$$

As $\mu_0\,\mathbf{u}$ is the current density of the proper mass, (207) is the continuity equation expressing that the proper mass in the system considered has neither sources nor sinks.

By means of (39) we can now define a four-velocity U_i at any point in the matter and at any time. As $U_i = U_i(x)$ is a function of the space-time coordinates, we thus have a four-vector field in (3+1)-space. In the same way, the invariant mass density μ^0 defined by (195) and (199) can be regarded as a scalar field in (3+1)-space, μ^0 being then a function of the space-time coordinates in any system of inertia:

$$\mu^0 = \mu_0(x)\sqrt{\{1-u^2(x)/c^2\}} = \mu^0(x). \tag{208}$$

By multiplication of the scalar μ^0/c and the vector U_ι we obtain a new four-vector

$$c_\iota = \frac{\mu^0 U_i}{c}, \tag{209}$$

which may be called the four-current density of proper mass. According to (39) and (198) we get for the components of c_i

$$c_\iota = \left\{\frac{\mu^0 \mathbf{u}}{\sqrt{(c^2-u^2)}}, \frac{i\mu^0}{\sqrt{(1-u^2/c^2)}}\right\} = \left\{\frac{\mu_0 \mathbf{u}}{c}, i\mu_0\right\}, \tag{210}$$

and the continuity equation (207) can be written in the tensor form

$$\frac{\partial c_\iota}{\partial x_i} = \frac{1}{c}\frac{\partial \mu^0 U_\iota}{\partial x_i} = 0. \tag{211}$$

The left-hand side of (211) is equal to the four-dimensional divergence (178) of the four-current density, and the covariance of this equation under rotations in (3+1)-space is therefore evident.

The forces acting on the different parts of continuously distributed matter are partly external impressed forces, partly elastic forces acting between neighbouring parts in the continuum. In this chapter, we shall completely neglect the last-mentioned forces and postpone the consideration of the elastic forces to Chapter VI. Here we therefore treat the matter as a kind of incoherent dust. The impressed forces, however, are assumed to be volume forces which in every system of inertia S can be described by a force density $\mathbf{f}$ so defined that $\mathbf{f}\,\delta V$ is equal to the force on the volume element δV.

Let us now consider the motion of a small material particle with the volume δV and with the proper mass $\mu_0\,\delta V = \mu^0\,\delta V^0$. When U_ι is the four-velocity, the four-momentum of this particle is, according to (50),

$$p_\iota = \mu^0\,\delta V^0 U_\iota, \tag{212}$$

and for the four-force (54) we obtain by means of (196)

$$F_\iota = \left(\frac{\mathbf{f}\,\delta V}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{f}.\mathbf{u})\,\delta V/c}{\sqrt{(1-u^2/c^2)}}\right) = \left(\mathbf{f}, \frac{i(\mathbf{f}.\mathbf{u})}{c}\right)\delta V^0. \tag{213}$$

Since δV^0 is an invariant and F_i is a four-vector, the quantity

$$f_i = \left(\mathbf{f}, \frac{i(\mathbf{f}.\mathbf{u})}{c}\right) \tag{214}$$

will also form a four-vector, the four-force density. The spatial components of f_ι are equal to the ordinary force density and f_4 is equal to the mechanical work done per unit of time and of volume multiplied by i/c.

The motion of the particle considered is now described by equation (55). Substituting from (212), (213) in (55) gives, when the proper mass is conserved, i.e. when $\frac{d}{d\tau}(\mu^0\,\delta V^0) = 0$,

$$\mu^0 \frac{dU_i}{d\tau} = f_i, \tag{215}$$

where f_i is given by (214). Since μ^0 and $d\tau$ are invariants, both sides of this equation represent four-vectors. The first three equations (215) are again the equations of motion, while the fourth equation expresses the theorem of conservation of energy. According to (214) and (39)

$$f_i\, U_i = 0, \tag{216}$$

which, as discussed in § 38, is essential for the conservation of the proper mass. By multiplying (215) by U_i and summing over i, the left-hand side becomes zero on account of (41′), and the equations (215) therefore appear to be compatible with (216).

The fundamental equations of mechanics assume the simple form (215) only in systems in which the proper mass is conserved. If we drop this assumption we have, in accordance with (59), to use the following expression for the four-force density:

$$f_i = \left\{\mathbf{f}, \frac{i}{c}((\mathbf{f}.\mathbf{u})+q)\right\}, \tag{217}$$

where q is the amount of non-mechanical energy developed per unit of volume and time, so that $(\mathbf{f}.\mathbf{u})+q$ represents the total effect per unit volume. In this case the left-hand side of (55) becomes

$$\frac{d}{d\tau}(\mu^0 U_i\,\delta V^0) = \frac{d(\mu^0 U_i)}{d\tau}\,\delta V^0 + \mu^0 U_i \frac{d\delta V^0}{d\tau}. \tag{218}$$

Since δV^0 is the volume of the particle and $d\tau$ is the time increase, both measured in the rest system S^0, we obtain by applying equation (204) in the rest system

$$\frac{d\delta V^0}{d\tau} = \delta V^0\,\mathrm{div}^0\,\mathbf{u}^0, \tag{219}$$

where div^0 means differentiations with respect to the space coordinates (x_i^0) in S^0. In the rest system $\mathbf{u}^0 = 0$, on the other hand, the derivatives of $\mathbf{u}^0$ with respect to the coordinates need not be zero. By means of the expressions (39) for U_i and U_i^0 we now get, since the quantity $\partial U_i/\partial x_i$ is invariant,

$$\frac{\partial U_k}{\partial x_k} = \frac{\partial U_k^0}{\partial x_k^0} = \mathrm{div}^0\,\mathbf{u}^0. \tag{220}$$

Using the equations (217)–(220) in (55) we then obtain in the general case, instead of (215),

$$\frac{d}{d\tau}(\mu^0 U_i)+\mu^0 U_i\frac{\partial U_k}{\partial x_k} = f_i, \tag{221}$$

where f_i is now given by (217). Since, furthermore,

$$\frac{d}{d\tau}(\mu^0 U_i) = \frac{\partial(\mu^0 U_i)}{\partial x_k}\frac{dx_k}{d\tau} = \frac{\partial(\mu^0 U_i)}{\partial x_k}U_k, \tag{222}$$

(221) can also be written

$$\frac{\partial}{\partial x_k}(\mu^0 U_i U_k) = f_i, \tag{223}$$

which thus represent the fundamental equations governing the motion of a continuous mass distribution in the general case in which the proper mass may not be conserved.

From (217) and (39) we now obtain

$$f_i U_i = -\frac{q}{\sqrt{(1-u^2/c^2)}} = -q^0, \tag{224}$$

where q^0 is the non-mechanical effect in the rest system. This quantity of course is an invariant, and we get from (224)

$$q = q^0\sqrt{(1-u^2/c^2)} \tag{225}$$

Multiplying this equation by the invariant

$$\delta V\Delta t = \delta V^0\Delta\tau$$

we again have the equation (66), since $\Delta Q = q\,\delta V\Delta t$ is the amount of heat conveyed to a certain material particle during the time Δt. If we multiply (223) by U_i and sum, we obtain, using (41), (41′), and (224),

$$\frac{\partial(\mu^0 U_k)}{\partial x_k} = \frac{q^0}{c^2}, \tag{226}$$

or, according to (209) and (210),

$$\frac{\partial\mu_0}{\partial t}+\operatorname{div}(\mu_0\,\mathbf{u}) = \frac{q^0}{c^2}. \tag{227}$$

These equations are thus generalizations of the equations (211) and (207) for the case where the system contains sources for the proper mass, and the expression q^0/c^2 for the source density is in accordance with Einstein's general theorem (III. 74).

In the preceding considerations it has been tacitly assumed that *no* heat conduction takes place in the matter, so that the transport of heat therefore occurs only by means of convection. In the more general case where heat conduction also occurs, it must be taken into account that the energy transported by means of heat conduction represents an extra, non-material momentum which would cause a change in the first three equations (223).

51. The kinetic energy-momentum tensor

The quantity appearing on the left-hand side of (223)

$$\theta_{ik} = \mu^0 U_i U_k \tag{228}$$

is a symmetrical tensor of rank 2 which is called the kinetic energy-momentum tensor.† (223) can thus be written

$$\frac{\partial \theta_{ik}}{\partial x_k} = f_i, \tag{229}$$

i e. the four-force density is equal to the divergence of the kinetic energy-momentum tensor. By means of (39), (198), and (199) we obtain the following expressions for the components of this tensor:

$$\theta_{44} = -\frac{\mu^0 c^2}{1-u^2/c^2} = -\frac{\mu_0 c^2}{\sqrt{(1-u^2/c^2)}} = -\mu c^2 = -h, \tag{230}$$

where h is the energy density or, more precisely, the sum of the kinetic energy and the proper energy per unit volume. Multiplication of (230) by δV makes the right-hand side, apart from the minus sign, equal to the energy of the small material particle inside the volume δV.

The three components θ_{4i} form the components of a spatial vector which can be written

$$(\theta_{41}, \theta_{42}, \theta_{43}) = \frac{i}{c}\frac{\mu^0 c^2}{1-u^2/c^2}\mathbf{u} = \frac{i}{c}\frac{\mu_0 c^2}{\sqrt{(1-u^2/c^2)}}\mathbf{u} = \frac{i}{c}\mu c^2\mathbf{u} = \frac{i}{c}h\mathbf{u}. \tag{231}$$

The three components θ_{i4} can in the same way be written

$$(\theta_{14}, \theta_{24}, \theta_{34}) = \frac{ic\mu_0\,\mathbf{u}}{\sqrt{(1-u^2/c^2)}} = ic\mathbf{g}. \tag{232}$$

In view of the equations (229) (cf. (236) and (238)) the quantities

$$\frac{c}{i}\theta_{4i} = h\mathbf{u} \quad \text{and} \quad \theta_{i4}/ic = \mathbf{g}$$

must be interpreted as *energy current density* and *momentum density*, respectively. The symmetry of the energy-momentum tensor which involves $\theta_{i4} = \theta_{4i}$ or

$$\mathbf{g} = \frac{h}{c^2}\mathbf{u} \tag{233}$$

then simply expresses that the energy h corresponds to a mass h/c^2.

† H. Minkowski, *Gott. Nachr.*, p. 53 (1908); *Math. Ann.* **68**, 472 (1910).

The spatial part $\theta_{\iota\kappa}$ of the energy-momentum tensor can be written

$$\theta_{\iota\kappa} = \frac{\mu_0 u_\iota u_\kappa}{\sqrt{(1-u^2/c^2)}} = g_\iota u_\kappa, \tag{234}$$

where g_ι and u_κ are the components of the space vectors $\mathbf{g}$ and $\mathbf{u}$.

Just as the quantities (231) represent the energy current density, each row of the space tensor $\theta_{\iota\kappa}$ in (234), e.g.

$$(\theta_{\iota 1}, \theta_{\iota 2}, \theta_{\iota 3}) = g_\iota \mathbf{u}, \tag{235}$$

can be interpreted as the current density of the momentum component g_ι. Therefore $\theta_{\iota\kappa}$ is also called the *momentum current tensor*.

The fourth equation (229), expressing the energy conservation, can now, by means of (230), (231), and (217), be written

$$\frac{\partial h}{\partial t} + \operatorname{div}(h\mathbf{u}) = (\mathbf{f}.\mathbf{u}) + q \tag{236}$$

or

$$\frac{\partial \mu}{\partial t} + \operatorname{div}(\mu\mathbf{u}) = \frac{(\mathbf{f}.\mathbf{u}) + q}{c^2}. \tag{237}$$

This equation is exactly analogous to equation (227). While q^0/c^2 expresses the source density of proper mass, we see that the source density for relativistic mass is $((\mathbf{f}.\mathbf{u})+q)/c^2$. Again this is in accordance with Einstein's relation (III. 74).

In the same way, the first three equations (229), the equations of motion, can be written

$$\frac{\partial g_\iota}{\partial t} + \frac{\partial \theta_{\iota\kappa}}{\partial x_\kappa} = f_\iota, \tag{238}$$

$$\frac{\partial g_\iota}{\partial t} + \operatorname{div}(g_\iota \mathbf{u}) = f_\iota, \tag{239}$$

where we have made use of (217), (232), and (234). Equation (239) plays the same role for the momentum as does the analogous equation (236) for the energy. It regulates the flow of momentum in the material continuum, the force density $\mathbf{f}$ appearing here as a source of momentum.

If we multiply (238) by x_κ and subtract the corresponding equation obtained by interchanging ι and κ, we get

$$\frac{\partial}{\partial t}(g_\iota x_\kappa - g_\kappa x_\iota) + \frac{\partial}{\partial x_\lambda}(\theta_{\iota\lambda} x_\kappa - \theta_{\kappa\lambda} x_\iota) - \theta_{\iota\kappa} + \theta_{\kappa\iota} = f_\iota x_\kappa - f_\kappa x_\iota. \tag{240}$$

On account of the symmetry of the tensor $\theta_{\iota\kappa}$, the last two terms on the left-hand side cancel each other. Introducing the angular momentum density and the density of momentum of force by

$$m_{\iota\kappa} = x_\iota g_\kappa - x_\kappa g_\iota, \tag{241}$$

$$d_{\iota\kappa} = x_\iota f_\kappa - x_\kappa f_\iota \tag{242}$$

(cf. equations (82)–(86)), (240) can be written

$$\frac{\partial m_{\iota\kappa}}{\partial t}+\frac{\partial}{\partial x_\lambda}(m_{\iota\kappa}u_\lambda) = d_{\iota\kappa}, \tag{243}$$

where we have applied the expression (234) for $\theta_{\iota\kappa}$. This equation can also be given the form

$$\frac{\partial m_{\iota\kappa}}{\partial t}+(\mathbf{u}.\mathrm{grad})m_{\iota\kappa}+m_{\iota\kappa}\operatorname{div}\mathbf{u} = d_{\iota\kappa}, \tag{244}$$

and, when we multiply (244) by δV, we obtain, by means of (200) and (204),

$$\frac{d}{dt}(m_{\iota\kappa}\delta V) = d_{\iota\kappa}\delta V. \tag{245}$$

This equation expresses the angular momentum theorem for a small material particle with the volume δV. We thus have seen that the symmetry of the kinetic energy-momentum tensor is a very important property, since the symmetry of the spatial part is essential for the validity of the angular momentum theorem in its usual form, while the equation $\theta_{\iota 4} = \theta_{4\iota}$, i e. (233), is an expression for Einstein's theorem of the inertia of energy.

The kinetic energy-momentum tensor satisfies the relation

$$\theta_{ik}U_k = \mu^0 U_i U_k U_k = -\mu^0 c^2 U_i = -h^0 U_i, \tag{246}$$

where h^0 is the energy density in the rest system.

V
ELECTRODYNAMICS IN THE VACUUM

52. The fundamental equations of electrodynamics in the vacuum. Four-current density for electric charge

In Chapter III we have seen that it is necessary to change the fundamental equations of mechanics in order to bring them into accordance with the principle of relativity. This is not so with the equations of electrodynamics in the vacuum, the Maxwell equations, which, as we shall see, are already covariant under Lorentz transformations.†

Let us imagine two teams of experimental physicists who have installed their laboratories in two different inertial systems S and S' and who independently are performing electromagnetic experiments. By means of electrically charged test bodies and magnetic compass needles the physicists in S will be able in the customary way to determine the electric field vector $\mathbf{E}$ and the magnetic field vector $\mathbf{H}$ as a function of the space-time coordinates $\mathbf{x}$ and t in S. By the same procedure the physicists in S' will be able to determine electric and magnetic field vectors $\mathbf{E}'$ and $\mathbf{H}'$ as functions of the coordinates $\mathbf{x}'$ and t' in S'. Furthermore, the two groups of physicists can, independently of each other, determine the charge densities ρ and ρ' in S and S'. In the present chapter we shall consider only electromagnetic phenomena in the vacuum; as neither conductors nor dielectric and magnetic substances are present here, the only type of electric currents occurring are convection currents. The current densities in S and S' will thus be $\rho\mathbf{u}$ and $\rho'\mathbf{u}'$, where $\mathbf{u}$ and $\mathbf{u}'$ are the velocities with which the charges move in S and S', respectively. All these quantities will be certain functions of the space and time coordinates in S and S'.

Now, according to the principle of relativity, the equations determining the fields as functions of the charge distribution should have the same form in S and S'. Consequently, both groups of physicists should as a result of their experiments be led to the Maxwell–Lorentz field equations for empty space. In S we thus have, applying Heaviside's units,

$$\operatorname{div}\mathbf{H} = 0, \qquad \operatorname{curl}\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = 0, \tag{1 a}$$

$$\operatorname{div}\mathbf{E} = \rho, \qquad \operatorname{curl}\mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{\rho\mathbf{u}}{c}, \tag{1 b}$$

† H. Poincaré, *C.R.* **140**, 1504 (1905), *Rend. Pal.* **21**, 129 (1906), A. Einstein, *Ann. d. Phys.* **17**, 891 (1905), H. Minkowski, see ref., Chap. IV, p. 136.

and in S' we get equations resulting from (1) by adding a prime to all quantities. The equations (1) are identical with the fundamental equations of Lorentz's classical electron theory.

While the connexion between $\mathbf{u}$ and $\mathbf{u}'$ is given by (II. 55), we do not yet know the connexion between ρ and ρ' or between $\mathbf{E}$ and $\mathbf{H}$, on the one hand, and $\mathbf{E}'$ and $\mathbf{H}'$ on the other. However, it is one of the most fundamental experiences that electric charge is conserved, a property which, in analogy to (IV. 207), can be expressed by the continuity equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{u}) = 0. \tag{2}$$

This equation is a simple consequence of Maxwell's equations (1 *b*). Obviously, a similar equation must be valid in S', viz.

$$\frac{\partial \rho'}{\partial t'} + \operatorname{div}'(\rho' \mathbf{u}') = 0. \tag{2'}$$

The connexion between ρ and ρ' must now be such that for arbitrary charge and current distributions (2′) is a consequence of (2). Defining four quantities in S by

$$s_i = \left(\frac{\rho \mathbf{u}}{c}, i\rho\right), \tag{3}$$

and analogous quantities s'_i in S', (2) and (2′) may be written

$$\frac{\partial s_i}{\partial x_i} = 0 \tag{4}$$

and

$$\frac{\partial s'_i}{\partial x'_i} = 0. \tag{4'}$$

In Appendix 2 it is shown that, if (4′) is to be a consequence of (4) for all possible charge and current distributions, the connexion between s_i and s'_i must be given by

$$s'_i = \alpha_{ik} s_k, \tag{5}$$

where α_{ik} are the coefficients in the coordinate transformation (IV. 3) connecting the systems S and S'. Thus, s_i is a four-vector, called the four-current density, and (4) expresses that the divergence of the four-current density is zero (cf. IV. 178).

If we multiply the invariant

$$s_i s_i = s'_i s'_i \tag{6}$$

by -1, we obtain with the help of (3) an invariant

$$\rho^2\left(1 - \frac{u^2}{c^2}\right) = \rho'^2\left(1 - \frac{u'^2}{c^2}\right) = \rho^{02}, \tag{7}$$

where ρ^0 is the charge density in the rest system S^0. Hence we get

$$\rho = \frac{\rho^0}{\sqrt{(1-u^2/c^2)}}. \tag{8}$$

By means of (8), the equation (3) can be written in the form

$$s_\iota = \frac{\rho^0 U_\iota}{c}, \tag{9}$$

where U_ι is the four-velocity defined by (IV. 39). Equation (9) is quite analogous to the expression (IV. 209) for the four-current density of the proper mass.

Let us now consider the charge $\rho\,\delta V$ connected with a material volume element δV. If δV^0 denotes the corresponding volume in the rest system, we have $\delta V = \delta V^0\sqrt{(1-u^2/c^2)}$, which together with (8) gives

$$\rho\,\delta V = \rho^0\,\delta V^0. \tag{10}$$

Hence, the electric charge of a certain material volume element is an invariant, and the same is therefore also true for the total charge of a material body. This important theorem of the invariance of electric charge is thus a consequence of the validity of the continuity equation in every system of inertia. It can also be made plausible by the following reasoning. Consider a charged particle of charge e originally at rest in S. Under the action of a force the particle is accelerated until it has the same velocity $\mathbf{v}$ as has S' relative to S. Since the charge of a particle is conserved during acceleration, the particle has still the charge e relative to S. On the other hand, the particle now has the velocity zero relative to S' and, since it is now in the same situation relative to S' as it previously was relative to S, the charge e' of the particle relative to S' must be assumed to be equal to the constant charge e relative to S. Therefore we must have at any time $e' = e$, in agreement with the equation (10).

53. Covariance of the fundamental equations of electrodynamics under Lorentz transformations. The electromagnetic field tensor

In every system of inertia S we now define a quantity $F_{\iota k}$ by the equations

$$F_{\iota k} = -F_{k\iota}, \qquad (F_{23}, F_{31}, F_{12}) = \mathbf{H}, \qquad (F_{41}, F_{42}, F_{43}) = i\mathbf{E}, \tag{11}$$

i.e.

$$F_{\iota k} = \begin{pmatrix} 0 & H_z & -H_y & -iE_x \\ -H_z & 0 & H_x & -iE_y \\ H_y & -H_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix}. \tag{12}$$

The equations (1 *a*) can then be written

$$\frac{\partial F_{ik}}{\partial x_l}+\frac{\partial F_{kl}}{\partial x_i}+\frac{\partial F_{li}}{\partial x_k}=0. \tag{13}$$

Since the expression on the left-hand side in (13) is completely antisymmetric in the three indices i, k, l, (13) represents only four independent equations which are obtained, for example, by putting (i,k,l) equal to (1, 2, 3), (4, 2, 3), (4, 3, 1), (4, 1, 2), respectively. It is easily verified that these four equations are identical with the four equations (1 *a*).

From the general validity of the equations (4) in every system of inertia we concluded that the quantities s_i are the components of a four-vector. In the same way we can conclude that the quantities F_{ik} must transform as the components of an antisymmetrical tensor if the equations (13), as should be required, are valid in every system of inertia. The tensor F_{ik} thus defined is called the electromagnetic field tensor, and the equations (13) or (1 *a*) express that the curl of this tensor is zero (cf. IV. 184).

On account of (IV. 187) the equations (13) may also be written

$$\operatorname{div}_i F^* = \frac{\partial F^*_{ik}}{\partial x_k} = 0, \tag{14}$$

where F^*_{ik} is the pseudo-tensor dual to F_{ik} obtained by the substitution $\mathbf{E}\to\mathbf{H}$, $\mathbf{H}\to-\mathbf{E}$ in the expression (12) for F_{ik}.

The connexion (11) between F_{ik} and $\mathbf{E}$, $\mathbf{H}$ is the same as in (IV. 80, 80′), which indicates that $\mathbf{E}$ and $\mathbf{H}$ behave as a polar and an axial vector, respectively, under pure spatial transformations. For a general Lorentz transformation without rotation we have the transformation equations (IV. 81′) for $\mathbf{E}$ and $\mathbf{H}$. These equations may also be written in the form

$$\left.\begin{aligned}\mathbf{E}' &= \frac{\mathbf{E}+\frac{\mathbf{v}}{v^2}(\mathbf{v}\cdot\mathbf{E})\left\{\sqrt{\left(1-\frac{v^2}{c^2}\right)}-1\right\}+\frac{1}{c}(\mathbf{v}\times\mathbf{H})}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}\\ \mathbf{H}' &= \frac{\mathbf{H}+\frac{\mathbf{v}}{v^2}(\mathbf{v}\cdot\mathbf{H})\left\{\sqrt{\left(1-\frac{v^2}{c^2}\right)}-1\right\}-\frac{1}{c}(\mathbf{v}\times\mathbf{E})}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}\end{aligned}\right\}. \tag{15}$$

The division of the field into an electric and a magnetic field, which is forced upon us by our measuring instruments, thus has no absolute meaning. If, for example, we have a purely electrostatic field in S, i.e. $\mathbf{H}=0$, there will, according to (15), be a magnetic field $\mathbf{H}'\neq 0$ in S'.

This is also quite clear from a physical point of view, since a purely electrostatic field in S' means that all charges are at rest relative to S. Relative to S' the charges will therefore move with the velocity $-\mathbf{v}$. Consequently we have in S' a stationary current which causes a magnetic field in S'.

By means of (3) and (12) the second set of Maxwell's equations (1 *b*) can now be written

$$\frac{\partial F_{ik}}{\partial x_k} = s_i \tag{16}$$

Since the left-hand side in (16) is the divergence of the electromagnetic field tensor (cf. IV. 183), the left-hand and the right-hand sides are transformed in the same way, viz. as a four-vector. Hence, the covariance of the equations (1 *b*) is a consequence of the covariance of equations (1 *a*) and of the continuity equation (4). This is a strong argument in favour of the exact validity of (1 *b*) and it is seen, in particular, that the term

$$\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t},$$

Maxwell's displacement current, is absolutely necessary for the covariance of the equations (1 *b*).

If we form the divergence of the vector equation (16) we get, on account of the antisymmetry of the electromagnetic field tensor,

$$\frac{\partial s_i}{\partial x_i} = \frac{\partial^2 F_{ik}}{\partial x_i \partial x_k} = 0, \tag{17}$$

i e. the continuity equation (4).

54. The four-potential. Gauge transformation

As is well known, **E** and **H** can, as a consequence of the equations (1 *a*), be written in the form

$$\mathbf{H} = \operatorname{curl}\mathbf{A}, \qquad \mathbf{E} = -\operatorname{grad}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{18}$$

where the vector potential **A** and the scalar potential ϕ can always be chosen in such a way that they satisfy the Lorentz condition

$$\operatorname{div}\mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0. \tag{19}$$

This can be accomplished in every system of inertia. If we now define four quantities A_i in every system of inertia by

$$A_i = (\mathbf{A}, i\phi), \tag{20}$$

(18) and (19) can be written

$$F_{ik} = \frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}, \tag{21}$$

$$\frac{\partial A_i}{\partial x_i} = 0. \tag{22}$$

Since F_{ik} is a tensor, it must be possible to choose the potentials in the different systems of inertia such that the A_i are transformed as the components of a four-vector, the four-potential. According to (21) and (22) the electromagnetic field tensor is equal to the curl of the four-potential which has a divergence equal to zero. Moreover, from (IV. 184–186) it follows that Maxwell's equations (13) are a consequence of (21).

When F_{ik} is given, the four-potential A_i is by no means uniquely determined by (21); for, if A_i satisfies (21), the functions

$$A_i^* = A_i + \frac{\partial \psi}{\partial x_i}, \tag{23}$$

where ψ is an arbitrary scalar, will also satisfy (21).

The transformation (23) is called a gauge transformation and the measurable quantities F_{ik} are invariant under such transformations. The Lorentz condition restricts the class of permitted gauge transformations, but there is still a great variety of potentials A_i satisfying this condition. Substitution from (23) in (22) leads to the condition

$$\frac{\partial^2 \psi}{\partial x_i \partial x_i} = \Box \psi = 0. \tag{24}$$

If ψ is an arbitrary solution of (24), A_i^* will thus satisfy both (21) and (22) if A_i satisfies these equations.

Inserting (21) in (16) we get, using (22),

$$\frac{\partial^2 A_i}{\partial x_k \partial x_k} = -s_i, \tag{25}$$

or

$$\Box A_i = -s_i. \tag{26}$$

Any solution of (26) which simultaneously satisfies (22) gives, by means of (21), a solution of Maxwell's equations (13) and (16).

55. Four-dimensional integral representation of the four-potential

The equations (25) have the form of usual potential equations in (3+1)-space; the solutions can therefore be found by a method exactly analogous to that applied in three dimensions.† We shall first write

† A. Sommerfeld, *Ann. d. Phys.* **33**, 649 (1910).

down the solution of the equations (25) on the assumption that all four coordinates x_i are real, so that the four-dimensional space is Euclidean. Let

$$R_\iota = x_\iota - x_\iota(P) \tag{27}$$

be a four-vector connecting a fixed point P with coordinates $x_\iota(P)$ and a variable point with coordinates x_ι. If $R^2 = (R_\iota R_\iota)$ denotes the square of the distance between these points a simple calculation shows that

$$\Box \frac{1}{R^2} = \frac{\partial^2}{\partial x_k \partial x_k} \frac{1}{R^2} = 0 \tag{28}$$

at any point $x_\iota \neq x_\iota(P)$. Further, if $\psi(x)$ and $\phi(x)$ are two arbitrary regular functions of $(x) = (x_\iota)$, we have

$$\phi \Box \psi - \psi \Box \phi = \frac{\partial}{\partial x_k}\left(\phi \frac{\partial \psi}{\partial x_k} - \psi \frac{\partial \phi}{\partial x_k}\right).$$

Now let us put $\psi = 1/R^2$ in this equation and integrate over the whole four-dimensional space which lies outside a (three-dimensional) sphere

$$R^2 = R_i R_\iota = (x_\iota - x_\iota(P))(x_\iota - x_\iota(P)) = a^2 \tag{29}$$

with radius a and with the point P as centre. If in the right-hand side we use $R_i = x_\iota - x_\iota(P)$ as integration variables instead of x_ι, we thus get, on account of (28),

$$-\int \frac{1}{R^2} \Box \phi \, dx_1 dx_2 dx_3 dx_4$$
$$= \int \frac{\partial}{\partial R_k}\left(\phi \frac{\partial}{\partial R_k} \frac{1}{R^2} - \frac{1}{R^2} \frac{\partial \phi}{\partial R_k}\right) dR_1 dR_2 dR_3 dR_4, \tag{30}$$

where the integration has to be extended over the four-dimensional region for which

$$R^2 = R_\iota R_\iota \geqslant a^2. \tag{31}$$

Since the integrand on the right-hand side of (30) is a sum of partial differential coefficients, the integral can be transformed into a three-dimensional integral over the sphere (29), provided that the function ϕ vanishes sufficiently rapidly at infinity. For the first term on the right-hand side of (30) we then obtain

$$\int \left[-\left(\phi \frac{\partial}{\partial R_1} \frac{1}{R^2} - \frac{1}{R^2} \frac{\partial \phi}{\partial R_1}\right)^+ + \left(\phi \frac{\partial}{\partial R_1} \frac{1}{R^2} - \frac{1}{R^2} \frac{\partial \phi}{\partial R_1}\right)^- \right] dR_2 dR_3 dR_4, \tag{32}$$

where the integration is to be extended over the sphere (29). Accordingly, $(\ )^+$ and $(\ )^-$ means that R_1 in the function in the brackets is to be put equal to the values

$$R_1 = \pm\{a^2 - (R_2^2 + R_3^2 + R_4^2)\}^{\frac{1}{2}} = \pm(a^2 - \rho_1^2)^{\frac{1}{2}} \tag{33}$$

and the domain of integration over the variables R_2, R_3, R_4 is defined by the inequality

$$\rho_1^2 = R_2^2 + R_3^2 + R_4^2 < a^2. \tag{34}$$

We get expressions similar to (32) for the three other terms corresponding to $k = 2, 3, 4$ on the right-hand side of (30); they are obtained from (32) by cyclic permutation of the indices (1, 2, 3, 4).

(The transformations just performed in (30) correspond to Green's theorem in three dimensions.)

Subsequently, letting $a \to 0$, the volume of the three-dimensional domain of integration in (32) will tend to zero as a^3 and, since

$$\frac{1}{R^2} = \frac{1}{a^2}$$

in the brackets $(\ )^{\pm}$, the second term inside the brackets will tend to zero. Since, moreover,

$$\frac{\partial}{\partial R_k}\frac{1}{R^2} = -\frac{2R_k}{R^4}, \tag{35}$$

we obtain for (32) in the limits of very small a, using (33),

$$2\phi(P)\int \frac{2\sqrt{(a^2-\rho_1^2)}}{a^4}\, dR_2\, dR_3\, dR_4, \tag{36}$$

where $\phi(P)$ is the value of the function ϕ at the point P. The domain of integration (34) is the interior of the sphere $\rho_1^2 = a^2$. The three other terms obtained by cyclic permutation of the indices 1, 2, 3, 4 are obviously equal to the first term. Therefore, introducing polar coordinates in the integration, we finally obtain for the right-hand side of (30)

$$16\phi(P)\frac{4\pi}{a^4}\int_0^a \sqrt{(a^2-\rho_1^2)}\rho_1^2\, d\rho_1 = 4\pi^2\phi(P), \tag{37}$$

and (30) becomes

$$4\pi^2\phi(P) = -\int \frac{1}{R^2}\Box\phi\, d^4x, \tag{38}$$

where d^4x is an abbreviation of $dx_1\, dx_2\, dx_3\, dx_4$. This equation holds for any regular function ϕ. If in particular ϕ is the function A_ι satisfying (26), we thus get the formula

$$4\pi^2 A_\iota(P) = \int \frac{s_\iota}{R^2}\, d^4x, \tag{39}$$

which allows us to calculate A_ι at an arbitrary point P in the four-dimensional space when s_ι is known at every point.

Until now we have assumed the variables x_ι to be real. However, in actual physical problems the four-current density s_ι is not given for real x_4, but only for purely imaginary x_4 values corresponding to $t < t(P)$.

In the complex x_4-plane s_ι is thus given only in the fat-faced part of the imaginary axis in Fig. 13. We therefore deform the original path of integration along the real axis into a loop L around this part of the imaginary axis, using the analytic continuation of the function s_ι in the integrand of (39). The expression (39) will then still be a solution of (26).

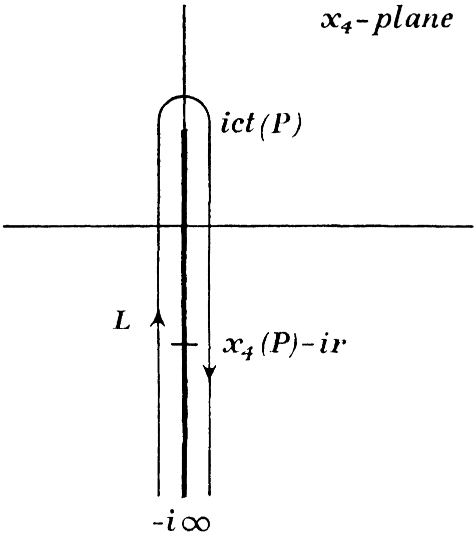

FIG. 13.

Since

$$\frac{\partial}{\partial x_\iota(P)} \frac{1}{R^2} = -\frac{\partial}{\partial x_\iota} \frac{1}{R^2}, \tag{40}$$

we get for the divergence of $A_\iota(P)$ in (39)

$$4\pi^2 \frac{\partial A_\iota(P)}{\partial x_\iota(P)} = -\int s_\iota \frac{\partial(1/R^2)}{\partial x_\iota} d^4x = \int \frac{\partial s_\iota}{\partial x_i} \frac{1}{R^2} d^4x = 0 \tag{41}$$

by partial integration and application of (4). The solution (39) thus also satisfies the Lorentz condition (22).

From (21), (39), (40), and (35) we thus obtain for the electromagnetic field tensor

$$2\pi^2 F_{ik}(P) = \int \frac{R_i s_k - R_k s_i}{R^4} d^4x. \tag{42}$$

56. Retarded potentials. Liénard–Wiechert's potentials for point charges

In performing the integrations in (39) the sequence of the integrations may be chosen arbitrarily and we shall first perform the integration over the x_4-variable along the path L in Fig. 13, keeping (x_1, x_2, x_3)

constant. If $r = |\mathbf{x}-\mathbf{x}(P)|$ denotes the spatial distance between the space points $\mathbf{x}$ and $\mathbf{x}(P)$ corresponding to the event points (x_i) and $(x_i(P))$, we have

$$\begin{aligned} R^2 &= r^2+(x_4-x_4(P))^2 \\ &= (x_4-x_4(P)+ir)(x_4-x_4(P)-ir). \end{aligned} \tag{43}$$

Therefore the quantity $1/R^2$ as a function of x_4 has a pole inside the loop L, viz. the point

$$x_4 = x_4(P)-ir \tag{44}$$

in the complex x_4-plane. Since the integrand in (39) has no other poles inside L, the path of integration can be deformed into a contour around the point (44) and, by means of Cauchy's theorem, we then obtain

$$\int\limits_L \frac{s_i}{R^2}\,dx_4 = -\frac{2\pi i s_i}{x_4-x_4(P)-ir}\bigg|_{x_4=x_4(P)-ir} = \pi\frac{s_i(\mathbf{x},t(P)-r/c)}{r}. \tag{45}$$

Consequently the potentials (39) assume the form

$$4\pi A_i(P) = \int \frac{s_i(\mathbf{x},t(P)-r/c)}{r}\,dV, \tag{46}$$

where the integration has to be extended over the usual three-dimensional space. The function s_i in the integrand is not to be taken at the time $t(P)$, but at the time $t(P)-r/c$, corresponding to the fact that all electromagnetic disturbances are propagated with the finite velocity c. Therefore the potentials (46) are called retarded potentials. If in (39) and (45) we had integrated along a curve obtained from L by reflection with respect to the point $x_4 = ict(P)$, we would have obtained another solution of (26) corresponding to advanced potentials. This solution, which connects the field at a certain point at a certain time with the future charge and current distributions, has, however, usually no immediate physical application.

Let us now consider the potential of a point charge e in arbitrary motion, the coordinates x_i of the point charge being given as functions of the time t:

$$x_i = x_i(t). \tag{47}$$

The four potentials can then be obtained from (46), but, since s_i in (46) on account of the retarded time variable has a rather complicated dependence on the variables of integration, it is easier in this case to go back to equation (39). Here we can now first perform the integration over the space coordinates and, since s_i is zero everywhere but at the points satisfying (47), we have, according to (3) and (IV. 39).

$$\int \frac{s_i}{R^2}\,dx_1\,dx_2\,dx_3 = \frac{1}{R^2}\left(\frac{e\mathbf{u}}{c}, ie\right) = \frac{e}{c}\left(1-\frac{u^2}{c^2}\right)^{\frac{1}{2}}\frac{U_i}{R^2}. \tag{48}$$

Here $\mathbf{u} = (dx_\iota/dt)$ is the velocity of the point charge given by (47), U_i is the corresponding four-velocity, and

$$R^2 = (x_\iota(t)-x_\iota(P))(x_\iota(t)-x_\iota(P))-c^2(t-t_p)^2 = r^2+(x_4-x_4(P))^2.$$

All these quantities are now functions of t or of the purely imaginary variable x_4, but by analytical continuation they can also be defined outside the imaginary axis in the x_4-plane. From (39) we thus get

$$4\pi^2 A_i(P) = \frac{e}{c}\int\limits_L \frac{\sqrt{(1-u^2/c^2)}U_i}{R^2}\,dx_4. \tag{49}$$

The integrand again has a pole at the point (44) in the complex plane, and near this x_4-value the denominator has the form

$$R^2 = \frac{dR^2}{dx_4}(x_4-x_4(P)+ir),$$

where the functions

$$\frac{dR^2}{dx_4} = 2R_k\frac{dR_k}{dx_4} = 2R_k U_k\frac{d\tau}{dx_4} = 2R_k U_k\frac{\sqrt{(1-u^2/c^2)}}{ic} \tag{50}$$

and $r(t)$ should be taken for the value of x_4 for which $x_4-x_4(P)+ir = 0$.

By means of Cauchy's theorem we thus obtain†

$$4\pi A_\iota(P) = \frac{eU_i}{U_k R_k}. \tag{51}$$

Here (R_k) denotes the four-vector leading from the fixed event point P to that point Q on the time track

$$x_\iota = x_i(\tau) \tag{52}$$

of the point charge in which the retrograde light cone

$$R^2 \equiv R_i R_i = 0 \tag{53}$$

originating from P intersects the time track. Also U_i should be taken at the point Q.

If $\mathbf{r}$ denotes the spatial part of the four-vector (R_ι) leading from P to Q we get for the denominator in (51)

$$U_i R_i = \frac{(\mathbf{u}.\mathbf{r})}{\sqrt{(1-u^2/c^2)}}+\frac{ic(x_4-x_4(P))}{\sqrt{(1-u^2/c^2)}} = \frac{(\mathbf{u}.\mathbf{r})+rc}{\sqrt{(1-u^2/c^2)}} \tag{54}$$

by means of (IV. 39) and (44). The equations (51) can thus be written

$$4\pi\mathbf{A}(P) = \frac{e\mathbf{u}/c}{r+(\mathbf{u}.\mathbf{r})/c}\bigg|_{t(P)-r/c}, \qquad 4\pi\phi(P) = \frac{e}{r+(\mathbf{u}.\mathbf{r})/c}\bigg|_{t(P)-r/c}. \tag{55}$$

† H. Minkowski, see ref., Chap. IV, p. 136.

The equations (55) represent Liénard–Wiechert's potentials of a moving point charge.

If we wish to calculate the electromagnetic field tensor $F_{ik}(P)$ at the event point P by means of (51) we must keep in mind that the proper time τ corresponding to the point Q is a function of the coordinates of the point P defined by the equation (53) or, according to (27) and (52), by

$$(x_i(\tau)-x_i(P))(x_i(\tau)-x_i(P)) = 0. \tag{56}$$

By differentiation of this equation with respect to $x_k(P)$ we obtain

$$R_i\left(\frac{dx_i}{d\tau}\frac{\partial\tau}{\partial x_k(P)}-\delta_{ik}\right) = 0$$

or

$$\frac{\partial\tau}{\partial x_k(P)} = \frac{R_k}{R_i U_i}. \tag{57}$$

From (21) and (51) we then get

$$4\pi F_{ik}(P) = \frac{d}{d\tau}\left(\frac{eU_k}{U_l R_l}\right)\frac{R_i}{U_m R_m}-\frac{d}{d\tau}\left(\frac{eU_i}{U_l R_l}\right)\frac{R_k}{U_m R_m},$$

or, using (IV. 41),

$$4\pi F_{ik} = \frac{e}{(U_l R_l)^2}\left(R_i\frac{dU_k}{d\tau}-R_k\frac{dU_i}{d\tau}\right)+$$
$$+\frac{e}{(U_l R_l)^3}\left(c^2-R_m\frac{dU_m}{d\tau}\right)(R_i U_k-R_k U_i). \tag{58}$$

Equation (58) could also be obtained directly from (42) if we first integrate over the space coordinates and subsequently use Cauchy's theorem in the following integration along the curve L in the x_4-plane. Here it should be remembered, however, that the function $1/R^4$ has a pole of *second* order at the point (44).

Since F_{ik} according to (58) has the form

$$4\pi F_{ik} = R_i a_k - R_k a_i, \tag{59}$$

where a_i is a four-vector, we have, according to (IV. 110, 111),

$$F_{ik} F^*_{ik} = 0, \tag{60}$$

where F^*_{ik} is the pseudo-tensor dual to F_{ik}, which is obtained from (12) by the substitution $(\mathbf{H},\mathbf{E}) \rightarrow (-\mathbf{E},\mathbf{H})$. The equation (60) is thus identical with the equation

$$(\mathbf{E}.\mathbf{H}) = 0. \tag{61}$$

The electric and magnetic field vectors are thus everywhere and in all systems of inertia perpendicular to each other. Further, we find by means of a simple calculation, applying (12), (58), and (53),

$$(4\pi)^2 F_{ik} F_{ik} = (4\pi)^2 2(|\mathbf{H}|^2 - |\mathbf{E}|^2) = -\frac{2e^2c^4}{(U_l R_l)^4}. \tag{62}$$

Consequently the invariant $|\mathbf{H}^2| - |\mathbf{E}^2|$ is always negative for the field of an arbitrarily moving point charge. This, together with (61), involves that for an arbitrary event point P we can always choose such a system of inertia that the components of the magnetic field vector in this system are zero at the point P. In order to obtain $\mathbf{H}' = 0$ in (15) one needs only to choose

$$\mathbf{v} = \frac{c(\mathbf{E} \times \mathbf{H})}{E^2} \tag{63}$$

and this is physically possible, since

$$v = \frac{cEH}{E^2} = \frac{cH}{E} < c \tag{64}$$

on account of (61) and (62).

57. The field of a uniformly moving point charge

Let us now in particular consider the field of a point charge moving with constant velocity. The time track of the charge is here a straight line in (3+1)-space with a direction defined by the constant four-velocity U_i. Since $dU_i/d\tau = 0$, (58) reduces to

$$4\pi F_{ik} = \frac{ec^2}{(U_l R_l)^3}(R_i U_k - R_k U_i). \tag{65}$$

By a suitable choice of the origin in the system of coordinates S we can always ensure that the x_4-axis and the time track are lying in the same plane In Fig. 14, which gives a representation of this plane, the line L represents the time track of the particle, and Q is the point at which the retrograde light cone from the arbitrary point P intersects the time track. Hence R_i is the vector leading from P to Q. If A is the projection on L of the point P the vector leading from A to P, with components $x_i^{(1)}$, is orthogonal to the four-velocity U_i, and $x_i^{(1)}$ is the projection of the vector $-R_i$ on a direction perpendicular to L; hence

$$x_i^{(1)} U_i = 0, \qquad -x_i^{(1)} R_i = x_i^{(1)} x_i^{(1)}. \tag{66}$$

Therefore we have

$$x_i^{(1)} = -R_i - \frac{U_i}{c^2}(U_l R_l). \tag{67}$$

It is immediately seen that the equation (67) satisfies both equations (66). From (67) and (53) we further obtain

$$x_i^{(1)} x_i^{(1)} = \frac{(U_i R_i)^2}{c^2}. \tag{67'}$$

In the rest system S' of the point charge the x'_4-axis is parallel to U_i and to the time track L. In this system we therefore have

$$x_i^{(1)'} = (\mathbf{r}', 0), \tag{68}$$

where $\mathbf{r}'$ is the space vector connecting the point charge with the space point p' corresponding to the event P. Since $U_i R_i$ is an invariant we get by means of (IV. 39), (27), and (44),

$$U_i R_i = U'_i R'_i = icR'_4 = cr' \quad (r' = |\mathbf{r}'|), \tag{69}$$

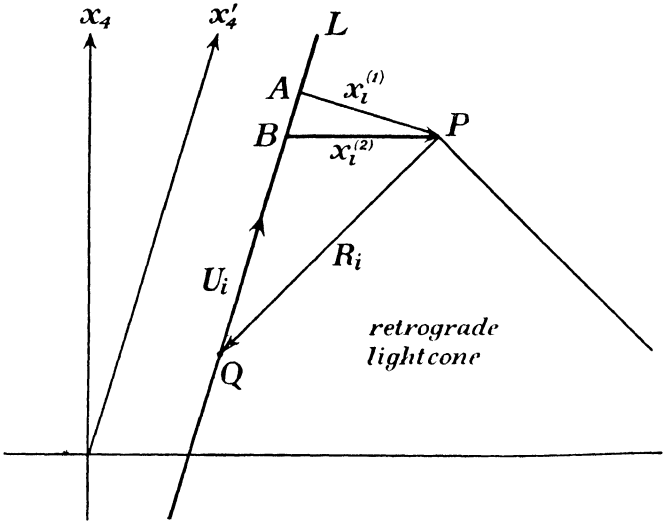

Fig. 14.

in accordance with (67′) and (68). Moreover, we get from (67)

$$R_i U_k - R_k U_i = -(x_i^{(1)} U_k - x_k^{(1)} U_i),$$

so that (65) can be written

$$4\pi F_{ik} = \frac{e}{cr'^3}(U_i x_k^{(1)} - U_k x_i^{(1)}). \tag{70}$$

In this expression we can obviously replace the vector $x_i^{(1)}$ by a vector leading from an *arbitrary* point on the curve L to the point P, since all such vectors have the form

$$x_i = x_i^{(1)} + aU_i, \tag{71}$$

where a is a constant. In an arbitrary system of coordinates S we now choose a vector $x_i^{(2)}$ such that the time component $x_4^{(2)}$ is zero. In Fig. 14 this vector is represented by the line BP. In S we thus have

$$x_i^{(2)} = (\mathbf{r}, 0), \tag{72}$$

where **r** is the space vector leading from the *simultaneous* position of the point charge to the point of observation p corresponding to the event P. (The **r** used here thus differs from the **r** used in the Liénard–Wiechert potentials which was the spatial part of the vector R_i.) We therefore have

$$4\pi F_{ik} = \frac{e}{cr'^3}(U_i x_k^{(2)} - U_k x_i^{(2)}), \tag{73}$$

which by means of (12), (72), and (IV. 39) leads to the following expressions for the electric and magnetic field vectors:

$$4\pi \mathbf{E} = \frac{e\mathbf{r}}{r'^3\sqrt{(1-u^2/c^2)}}, \qquad 4\pi \mathbf{H} = \frac{e}{cr'^3}\frac{\mathbf{u}\times\mathbf{r}}{\sqrt{(1-u^2/c^2)}}. \tag{74}$$

From Fig. 14 we see that the two vectors $x_i^{(2)}$ and $x_i^{(1)}$ have the same space components $\mathbf{r}'$ in S'. The connexion between $\mathbf{r}'$ and $\mathbf{r}$ is therefore obtained simply by using the reciprocal of the transformation equation (IV. 29) for the vector with components

$$x_i^{(2)} = (\mathbf{r}, 0), \qquad x_i'^{(2)} = (\mathbf{r}', x_4^{(2)'}).$$

Since the velocity of S' relative to S is **u**, we get

$$\mathbf{r}' = \mathbf{r} + \frac{\mathbf{u}}{u^2}\frac{(\mathbf{u}.\mathbf{r})(1-\sqrt{(1-u^2/c^2)})}{\sqrt{(1-u^2/c^2)}}, \tag{75}$$

hence

$$r' = \left\{r^2 + \frac{(\mathbf{u}.\mathbf{r})^2}{c^2-u^2}\right\}^{\frac{1}{2}}. \tag{76}$$

Decomposing the vectors **r** and $\mathbf{r}'$ into components perpendicular and parallel to **u**, respectively, (75) can be written

$$r'_{\parallel} = \frac{r_{\parallel}}{\sqrt{(1-u^2/c^2)}}, \qquad r'_{\perp} = r_{\perp}. \tag{77}$$

Hence the connexion between **r** and $\mathbf{r}'$ corresponds to a Lorentz contraction in the direction **u** (cf. II. 35).

The electric field vector lies in the direction of the radius vector **r**, while **H** is perpendicular both to **r** and **u**. The equations (74) are easily seen to be in agreement with the equations (61) and (62). The surfaces with constant values of the quantity (E^2-H^2) are obviously rotation ellipsoids, so-called Heaviside ellipsoids, which are obtained from spheres $r' =$ constant by a Lorentz contraction in the direction of motion of the point charge.

The expressions (74) may also be obtained in a much simpler way by means of the transformation equations (IV. 81′) for the electromagnetic field vectors. If S' denotes the rest system of the point charge, the electric field is spherically symmetrical and the magnetic field is zero in S', i.e.

$$4\pi \mathbf{E}' = \frac{e\mathbf{r}'}{r'^3}, \qquad \mathbf{H}' = 0. \tag{78}$$

Hence, from (IV. 81′) and (78),

$$4\pi \mathbf{E} = \frac{e}{r'^3} \frac{\mathbf{r}' + \frac{\mathbf{u}}{u^2}(\mathbf{u}\,.\,\mathbf{r}')\left\{\sqrt{\left(1-\frac{u^2}{c^2}\right)}-1\right\}}{\sqrt{\left(1-\frac{u^2}{c^2}\right)}},$$

$$4\pi \mathbf{H} = \frac{e}{c} \frac{(\mathbf{u}\times\mathbf{r}')}{r'^3\sqrt{(1-u^2/c^2)}}.$$

Since the equations reciprocal to (75) are

$$\mathbf{r} = \mathbf{r}' + \mathbf{u}\frac{(\mathbf{r}'\,.\,\mathbf{u})}{u^2}\left\{\sqrt{\left(1-\frac{u^2}{c^2}\right)}-1\right\} \tag{75′}$$

(cf. II. 35′), we again obtain the expression (74) for the electromagnetic field of a point charge in uniform motion.

From the latter deduction it results that the equations (74) are valid also everywhere outside a charged sphere provided that the charge distribution in the rest system is spherically symmetrical In this case $\mathbf{r}$ denotes the distance from the centre of the sphere while e is the total charge on the sphere.

In the same way we can also determine the field of a 'uniformly accelerated' point charge, i.e. a point charge performing a hyperbolic motion (cf. § 29). Again we can use two different methods· either we can directly apply the equation (58) which is valid for an arbitrarily moving charge, or we can introduce a system of coordinates S^* which follows the charge in its motion and solve Maxwell's equations in this system, and then transform back to the system S. However, the system of co-ordinates S^* will not be a system of inertia and, consequently, the application of this latter method requires the development of a general theory of relativity which allows the use of arbitrarily moving systems of coordinates. The necessary tools for an application of the second method will be provided in §§ 97 and 115.

58. The electromagnetic forces acting on charged matter

By means of the equations deduced in the preceding sections we are able to calculate the field created by an arbitrary charge and current distribution. Now we shall consider the opposite problem of the influence of a given field F_{ik} on the motion of electrically charged matter. Our first task will be to determine the force on an electrically charged particle of charge e moving in a given electromagnetic field with the velocity $\mathbf{u}$ relative to a certain system of inertia S. In accordance with the method outlined in § 29 we shall now introduce the inertial system S^0 in which the particle is momentarily at rest. In this system the force $\mathbf{F}^0$ must be

$$\mathbf{F}^0 = e\mathbf{E}^0, \tag{79}$$

by the very definition of the electric field vector $\mathbf{E}^0$ in this system. Introducing the equation (79) into (III. 43), and keeping in mind that the velocity of S^0 relative to S is equal to the particle velocity $\mathbf{u}$, one easily finds by means of the transformation equations (15) for the electromagnetic field vectors, Lorentz's expression for the force $\mathbf{F}$ in the system S, viz.

$$\mathbf{F} = e\left[\mathbf{E}+\frac{1}{c}(\mathbf{u}\times\mathbf{H})\right]. \tag{80}$$

This expression thus follows without extra hypotheses from the principle of relativity.

This deduction becomes much simpler if we use the four-dimensional representation and try to determine the expression for Minkowski's four-force. In the case considered, the proper mass is conserved and the four-force is thus defined by (IV. 54); therefore we have in the rest system

$$F_i^0 = (\mathbf{F}^0, 0). \tag{81}$$

If U_i is the four-velocity of the particle we can now form a four-vector with the components

$$\frac{e}{c} F_{ik}\, U_k. \tag{82}$$

In the rest system S^0 we have

$$U_i^0 = (0, 0, 0, ic), \tag{83}$$

and the components of the four-vector (82) in S^0 thus become, according to (12),

$$\frac{e}{c} F_{ik}^0\, U_k^0 = (e\mathbf{E}^0, 0). \tag{84}$$

From (79), (84), and (81) we see that the components of the four-vectors F_i and (82) are equal in the rest system, but two four-vectors whose components are equal in one system of coordinates are altogether identical. Therefore we must have

$$F_i = \frac{e}{c} F_{ik}\, U_k \tag{85}$$

in every system of coordinates.

Since $F_{ik} = -F_{ki}$ we have

$$F_i U_i = \frac{e}{c} F_{ik} U_k U_i \equiv 0 \tag{86}$$

in agreement with (IV. 57), and the equations of motion of the particle are given by (IV. 56) and (85)

$$m_0 \frac{dU_i}{d\tau} = \frac{e}{c} F_{ik} U_k. \tag{87}$$

If we calculate the components of F_i in (85) by means of (12) and (IV. 39), we get the equations (IV. 54) with **F** given by the Lorentz equation (80).

Now consider a continuous distribution of charged matter, with a four-current density (3), (9), i.e.

$$s_i = \left(\frac{\rho \mathbf{u}}{c}, i\rho\right) = \frac{\rho^0 U_i}{c}, \tag{88}$$

in a given external field. The problem is then to find the expression for the four-force density f_i defined by (IV. 214). This question can be solved on similar lines as above.

Consider a definite point in space at a definite time; the charged matter at this point is moving with a certain velocity. Now, let S^0 be the momentary rest system of the matter at this point. The components of s_i in this system are then $s_i^0 = (0, 0, 0, i\rho^0)$ and the four-vector

$$F_{ik} s_k \tag{89}$$

has the components

$$F_{ik}^0 s_k^0 = (\rho^0 \mathbf{E}^0, 0) \tag{90}$$

in S^0. These components are equal to the corresponding components (IV. 214) of the four-force in S^0

$$f_i^0 = (\mathbf{f}^0, 0); \tag{91}$$

for the force-density in the rest system must be given by

$$\mathbf{f}^0 = \rho^0 \mathbf{E}^0 \tag{92}$$

by the definition of the electric field vector.

The four-force density must therefore be equal to the four-vector (89) in every system of coordinates, hence

$$f_i = F_{ik} s_k. \tag{93}$$

From (IV. 214), (93), (88), and (12) we immediately get Lorentz's expression for the force density

$$\mathbf{f} = \rho\left[\mathbf{E} + \frac{1}{c}(\mathbf{u} \times \mathbf{H})\right]. \tag{94}$$

Further, we get, in view of the antisymmetry of the field tensor

$$f_\iota U_i = F_{\iota k} s_k U_i = \frac{\rho^0}{c} U_i F_{\iota k} U_k \equiv 0, \tag{95}$$

which, as discussed in § 50, means that the proper mass is conserved in this case. Therefore the equations of motion have the simple form (IV. 215).

$$\mu^0 \frac{dU_i}{d\tau} = f_i. \tag{96}$$

59. Variational principle of electrodynamics

Maxwell's field equations and the equations of motion (96) may be derived from a certain variational principle formulated by Weyl and Born.† If we put

$$F_{ik} = \frac{\partial A_k}{\partial x_\iota} - \frac{\partial A_i}{\partial x_k},$$

where A_i is the four-potential, the first set of Maxwell's equations (13) is identically satisfied. Now, consider the invariant integral

$$L = \int_\Omega \mathscr{L}\, d\Sigma \tag{97}$$

with

$$\mathscr{L} = -\tfrac{1}{4} F_{ik} F_{\iota k} + A_\iota s_\iota - \mu^0 c^2 = -\frac{1}{4}\left(\frac{\partial A_k}{\partial x_\iota} - \frac{\partial A_i}{\partial x_k}\right)\left(\frac{\partial A_k}{\partial x_i} - \frac{\partial A_\iota}{\partial x_k}\right) + A_\iota s_\iota - \mu^0 c^2 \tag{98}$$

integrated over a certain domain Ω in (3+1)-space. s_i is the four-current density and μ^0 is the invariant mass density in the rest system.

In the first place, we shall now consider an arbitrary variation δA_i of the functions $A_i(x)$ for which $\delta A_\iota = 0$ at the boundary of the domain Ω, s_i and μ^0 being kept constant by this variation. For the variation of L we then get by partial integration, remembering that $\delta \dfrac{\partial A_i}{\partial x_k} = \dfrac{\partial \delta A_i}{\partial x_k}$,

$$\delta L = \int \delta\mathscr{L}\, d\Sigma = \int \left(F_{\iota k} \frac{\partial \delta A_i}{\partial x_k} + \delta A_i\, s_\iota\right) d\Sigma = \int \left(-\frac{\partial F_{ik}}{\partial x_k} + s_\iota\right) \delta A_\iota\, d\Sigma. \tag{99}$$

The condition

$$\delta L = 0 \tag{100}$$

for any variation of the kind considered then leads to the equations

$$-\frac{\partial F_{\iota k}}{\partial x_k} + s_i = 0,$$

i.e. to the second set of Maxwell's equations (16).

Next we shall consider a variation in which the A_ι are kept constant

† H. Weyl, *Raum-Zeit-Materie*, Berlin, 1918. See also M. Born, *Ann. d. Phys.* **28**, 571 (1909).

while the time tracks of the matter are varied. This means that only the last two terms in (98) are affected by the variation. Let us fix our attention on a definite infinitesimal piece of matter with volume dV and consider the infinitely thin tube of time tracks of this element of matter. If τ is the proper time, and $x_\iota = x_\iota(\tau)$ the space-time coordinates of this material particle, we have

$$U_\iota = \frac{dx_\iota}{d\tau}, \qquad s_\iota = \frac{\rho^0 U_\iota}{c} = \frac{\rho^0}{c}\frac{dx_\iota}{d\tau}.$$

We now choose the domain Ω in (97) as that part of this tube for which $\tau_1 < \tau < \tau_2$ and consider an arbitrary variation of the time track satisfying the condition $\delta x_\iota = 0$ for $\tau = \tau_1$ and $\tau = \tau_2$. Since

$$d\Sigma = \frac{1}{i}\, dV dx_4 = c\, dV^0 d\tau,$$

where dV^0 is the rest volume of the particle, we have

$$\int (A_\iota s_\iota - \mu^0 c^2)\, d\Sigma = \int \left(A_\iota \rho^0\, dV^0 \frac{dx_\iota}{d\tau} - \mu^0\, dV^0 c^3\right) d\tau$$

$$= de \int_{\tau_1}^{\tau_2} A_\iota \frac{dx_\iota}{d\tau}\, d\tau - c^3\, dm^0 \int_{\tau_1}^{\tau_2} d\tau. \tag{101}$$

Here $dm^0 = \mu^0\, dV^0$ and $de = \rho^0\, dV^0$ represent the total rest mass and total charge of the particle, respectively, quantities which are constant along the tube.

From the definition of $d\tau$ we get for the variation of the second integral by a variation $\delta x_\iota(\tau)$ of the kind considered

$$\delta \int_{\tau_1}^{\tau_2} d\tau = \frac{1}{c^2} \int_{\tau_1}^{\tau_2} \frac{dU_\iota}{d\tau}\, \delta x_\iota\, d\tau \tag{102}$$

(see § 91, p. 244).

Further, using

$$\delta \frac{dx_\iota}{d\tau} = \frac{d}{d\tau}\, \delta x_\iota, \qquad \frac{dA_\iota}{d\tau} = \frac{\partial A_\iota}{\partial x_k} \cdot \frac{dx_k}{d\tau} = \frac{\partial A_\iota}{\partial x_k} U_k,$$

we get

$$\delta \int_{\tau_1}^{\tau_2} A_i \frac{dx_\iota}{d\tau}\, d\tau = \int \left(\frac{\partial A_i}{\partial x_k}\, \delta x_k \frac{dx_i}{d\tau} + A_\iota \frac{d\delta x_\iota}{d\tau}\right) d\tau$$

$$= \int \left(\frac{\partial A_k}{\partial x_i} U_k - \frac{dA_\iota}{d\tau}\right) \delta x_\iota\, d\tau = \int \left(\frac{\partial A_k}{\partial x_\iota} - \frac{\partial A_\iota}{\partial x_k}\right) U_k\, \delta x_\iota\, d\tau$$

$$= \int_{\tau_1}^{\tau_2} F_{\iota k}\, U_k\, \delta x_i\, d\tau. \tag{103}$$

From (98), (101), (102), and (103) we thus get for the variation of L in this case

$$\delta L = c\int\left(\frac{\rho^0}{c}F_{\iota k}\,U_k-\mu^0\,\frac{dU_\iota}{d\tau}\right)\delta x_\iota\,dV_0\,d\tau$$

$$= c\int\left(f_\iota-\mu^0\,\frac{dU_\iota}{d\tau}\right)\delta x_\iota\,dV_0\,d\tau. \tag{104}$$

Thus the condition $\delta L = 0$ for any variation of the kind considered leads to the equation

$$\mu^0\,\frac{dU_\iota}{d\tau} = f_\iota,$$

i.e. just the equations of motion (96) of charged incoherent matter.

60. The electromagnetic energy-momentum tensor

We shall now show that by virtue of Maxwell's equations the four-force density (93) can be written in the form of the divergence of a symmetrical tensor. Introducing (16) into the expression (93) we get

$$f_\iota = F_{\iota l}\,s_l = F_{\iota l}\,\frac{\partial F_{lk}}{\partial x_k} = \frac{\partial(F_{\iota l}\,F_{lk})}{\partial x_k}-\frac{\partial F_{\iota l}}{\partial x_k}F_{lk}.$$

Further, we have

$$\frac{\partial F_{\iota l}}{\partial x_k}F_{lk} = \frac{\partial F_{k\iota}}{\partial x_l}F_{lk} = \frac{1}{2}\left(\frac{\partial F_{\iota l}}{\partial x_k}+\frac{\partial F_{k\iota}}{\partial x_l}\right)F_{lk} = -\frac{1}{2}\,\frac{\partial F_{lk}}{\partial x_i}F_{lk} = -\frac{1}{4}\,\frac{\partial}{\partial x_i}(F_{lk}\,F_{lk}),$$

where we have made use of (13) and of the antisymmetry of $F_{\iota k}$. Thus we may write f_i in the form

$$f_\iota = -\frac{\partial S_{\iota k}}{\partial x_k}, \tag{105}$$

with

$$S_{ik} = F_{\iota l}\,F_{kl}-\delta_{\iota k}\,\tfrac{1}{4}(F_{lm}\,F_{lm}). \tag{106}$$

According to the rules of tensor calculus, $S_{\iota k}$ is a tensor, it is symmetric,

$$S_{\iota k} = S_{k\iota}, \tag{107}$$

and it satisfies the identity

$$S_{\iota\iota} = F_{\iota l}\,F_{\iota l}-4\,.\,\tfrac{1}{4}(F_{lm}\,F_{lm}) \equiv 0. \tag{108}$$

Using (11) in (106) we get by a simple calculation the following expressions for the components of this tensor in terms of the electric and magnetic field vectors:

$$S_{\iota\kappa} = -t_{\iota\kappa}, \tag{109}$$

where

$$t_{\iota\kappa} = E_\iota\,E_\kappa+H_\iota\,H_\kappa-\tfrac{1}{2}(E^2+H^2)\delta_{\iota\kappa} \tag{110}$$

is Maxwell's stress tensor; further,

$$S_{\iota 4} = S_{4\iota} = \frac{i}{c}S_\iota, \tag{111}$$

where the vector $\mathbf{S}$ with components S_ι is Poynting's vector, i.e.

$$\mathbf{S} = c(\mathbf{E}\times\mathbf{H}). \tag{112}$$

Finally,

$$S_{44} = -W, \tag{113}$$

where

$$W = \tfrac{1}{2}(E^2+H^2) \tag{114}$$

is the electromagnetic field energy density.

The equation (105) with $i = 4$ may then be written

$$f_4 = \frac{i}{c}(\mathbf{f}.\mathbf{u}) = -\frac{i}{c}\operatorname{div}\mathbf{S}+\frac{1}{ic}\frac{\partial W}{\partial t},$$

$$(\mathbf{f}.\mathbf{u})+\operatorname{div}\mathbf{S}+\frac{\partial W}{\partial t} = 0, \tag{115}$$

which expresses the energy conservation law if W and $\mathbf{S}$ are interpreted as field energy density and current density, respectively. If we integrate (115) over a finite volume V in space enclosed by a fixed surface σ we get, by means of Gauss's theorem,

$$-\frac{d}{dt}\int\limits_V W\,dV = \int\limits_\sigma S_n\,d\sigma+\int\limits_V (\mathbf{f}.\mathbf{u})\,dV, \tag{116}$$

where S_n is the component of $\mathbf{S}$ along the outward normal of the surface element $d\sigma$. The decrease in field energy inside V per unit time is thus equal to the outward flux of field energy through the surface σ plus the total work done on the matter inside V by the electromagnetic forces.

For $i = 1,2,3$ (105) gives

$$f_\iota = \frac{\partial t_{\iota\kappa}}{\partial x_\kappa}-\frac{\partial}{\partial t}\left(\frac{S_\iota}{c^2}\right). \tag{117}$$

In the case of a static field the last term is zero and (117) is then exactly Maxwell's expression for the force density in a substance with $\epsilon = \mu = 1$. The present deduction is due to Minkowski.† As pointed out by Abraham,‡ the vector

$$\mathfrak{g} = \frac{\mathbf{S}}{c^2} \tag{118}$$

must be interpreted as electromagnetic momentum density if we want to have conservation of momentum for a closed system. For if we integrate (117) over the interior of a closed surface σ which contains the whole system we get

$$\int \mathbf{f}\,dV = -\frac{d}{dt}\int \mathfrak{g}\,dV \tag{119}$$

† H. Minkowski, see ref., Chap. IV, p. 136.

‡ M. Abraham, *Ann. d. Phys.* **10**, 105 (1903); Abraham–Becker, *Theorie der Elektrizitat*, vol. II, 6th ed., Leipzig, 1933.

since $\int \frac{\partial t_{\iota\kappa}}{\partial x_\kappa} dV$ by partial integrations can be transformed into an integral over the surface σ where the field and therefore also $t_{\iota\kappa}$ is zero. The left-hand side of (119) represents the total force exerted on the matter and is equal to the increase $d\mathbf{G}_m/dt$ per unit time of the mechanical momentum. Thus (119) may be written

$$\frac{d}{dt}\Big(\mathbf{G}_m + \int \mathfrak{g}\, dV\Big) = 0.$$

In order to obtain a constant total momentum we have thus to assume an electromagnetic momentum $\int \mathfrak{g}\, dv$ besides the mechanical momentum. It is true that from this argument we can only conclude that the electromagnetic momentum density must be $\mathbf{S}/c^2 + C$, where C is a constant, but since $\mathfrak{g}$ must vanish simultaneously with the field the constant C must be zero.

Now defining the velocity of propagation $\mathbf{w}$ of the electromagnetic energy by the equation

$$\mathbf{w} = \frac{\mathbf{S}}{W}, \tag{120}$$

we can write for the electromagnetic momentum density

$$\mathfrak{g} = \frac{\mathbf{S}}{c^2} = \frac{W}{c^2}\mathbf{w} \tag{121}$$

in analogy with (IV. 233), holding for the mechanical momentum density. This analogy is close only in the case where the velocity w defined by (120) satisfies the condition $w \leqslant c$. For a plane polarized electromagnetic wave we have $(\mathbf{E}.\mathbf{H}) = 0$ and $E = H$; thus we get

$$w = \frac{|S|}{W} = \frac{cEH}{\frac{1}{2}(E^2+H^2)} = c,$$

i.e. the energy in such a wave travels with the velocity of light.

For the field of a charged particle in arbitrary motion we have, on account of (61) and (62),

$$(\mathbf{E}.\mathbf{H}) = 0 \quad \text{and} \quad E/H = \alpha \geqslant 1,$$

i.e.

$$w = \frac{cEH}{\frac{1}{2}(E^2+H^2)} = \frac{2c\alpha}{1+\alpha^2} \leqslant c.$$

61. The total energy-momentum tensor

If we use the expressions (105) for the electromagnetic four-force density in the equations of motion (IV. 229), the laws of conservation of energy and momentum for the total system of matter and electromagnetic field take the form

$$\frac{\partial T_{ik}}{\partial x_k} = 0, \tag{122}$$

where $$T_{ik} = \theta_{ik} + S_{ik} \tag{123}$$

is the total energy-momentum tensor of the system. The components of this tensor are obtained from the components of θ_{ik} and S_{ik} given in §§ 51 and 60. We have

$$T_{\iota\kappa} = \frac{\mu_0 u_\iota u_\kappa}{\sqrt{(1-u^2/c^2)}} - t_{\iota\kappa}, \tag{124}$$

where $t_{\iota\kappa}$ is Maxwell's stress tensor.

$$T_{4\iota} = \frac{i}{c} S_\iota, \tag{125}$$

where $$\mathbf{S} = \frac{\mu_0 c^2}{\sqrt{(1-u^2/c^2)}} \mathbf{u} + c(\mathbf{E} \times \mathbf{H})$$

is the total energy current density. Further,

$$T_{\iota 4} = ic\mathbf{g}_\iota, \tag{126}$$

where $$\mathbf{g} = \frac{\mu_0 \mathbf{u}}{\sqrt{(1-u^2/c^2)}} + \frac{1}{c}(\mathbf{E} \times \mathbf{H})$$

is the total momentum density of the system.

Finally we have $$T_{44} = -h, \tag{127}$$

where $$h = \frac{\mu_0 c^2}{\sqrt{(1-u^2/c^2)}} + \tfrac{1}{2}(E^2 + H^2)$$

is the total energy density.

The sum of the diagonal components of this tensor is, on account of (108) and (IV. 198),

$$T_{ii} = \theta_{ii} = -\mu_0 c^2 \sqrt{(1-u^2/c^2)} = -\mu^0 c^2. \tag{128}$$

VI

GENERAL CLOSED SYSTEMS. MECHANICS OF ELASTIC CONTINUA. FIELD THEORY

62. Definition of a closed system

In Chapter V we have treated the case of incoherent charged matter under the influence of electromagnetic forces. We saw that the four-vector f_i describing these forces could be written as the four-dimensional divergence of a tensor which itself is a function of the field variables describing the electromagnetic field. As the principle of relativity requires that all signals are propagated with a velocity smaller than or equal to c, it is impossible to maintain the Newtonian idea of forces acting instantaneously over finite distances in space It seems necessary to assume that all forces between material bodies are transmitted by means of an intermediary field in a way similar to the case of electromagnetic forces. It is therefore generally assumed that all types of forces can be described by a four-force density f_i which is the divergence of a certain tensor S_{ik} depending on the field variables of the intermediary fields by analogy with (V. 105). For the total system of matter and fields we then get, in the same way as in § 61, the laws of conservation of energy and momentum in the form

$$\frac{\partial T_{ik}}{\partial x_k} = 0, \tag{1}$$

where T_{ik} is the total energy-momentum tensor of the closed system. The physical meaning of the components $T_{\iota 4}$ and $T_{4\iota}$ is the same as in (V 125, 126, 127), i.e.

$$T_{4\iota} = \frac{i}{c} S_\iota, \tag{2}$$

where $\mathbf{S}$ is the energy current density,

$$T_{\iota 4} = icg_\iota, \qquad T_{44} = -h, \tag{3}$$

where $\mathbf{g}$ and h represent the total momentum density and energy density, respectively. This should hold for any closed physical system, and hence also for elastic bodies, if the elastic stresses and energies are included.

The equations (1) for $i = 1, 2, 3$ may now be written

$$\frac{\partial T_{\iota\kappa}}{\partial x_\kappa} + \frac{\partial g_\iota}{\partial t} = 0, \tag{4}$$

which represents the law of conservation of momentum in differential form. $T_{\iota\kappa}$ is called the stress tensor or the momentum current tensor (cf § 51, p. 137).

Similarly, the equation (1) with $i = 4$ represents the continuity equation for energy.

$$\operatorname{div} \mathbf{S} + \frac{\partial h}{\partial t} = 0. \tag{5}$$

The total energy-momentum tensor of a closed system must be symmetric, i.e.

$$T_{ik} = T_{ki}. \tag{6}$$

The spatial part of this equation, i.e $T_{\iota\kappa} = T_{\kappa\iota}$, is essential for the validity of the conservation law of angular momentum (see §§ 51 and 63), and if this law is to hold in every system of inertia, the equation (6) must also be valid for the space-time components, i e. we must also have

$$T_{\iota 4} = T_{4\iota}, \tag{7}$$

or, on account of (2) and (3),

$$\mathfrak{g} = \mathbf{S}/c^2 \tag{8}$$

Defining the velocity of propagation $\mathbf{u}^*$ of the energy by

$$\mathbf{u}^* = \mathbf{S}/h \tag{9}$$

as in the case of the electromagnetic energy (V 120), the equation (7) or (8) can be written

$$\mathfrak{g} = \frac{h}{c^2}\,\mathbf{u}^*, \tag{10}$$

which is formally analogous to the equation (IV. 233) for the mechanical momentum density and thus shows that the energy density h corresponds to a mass density

$$\mu = h/c^2. \tag{11}$$

It should be remarked, however, that the velocity $\mathbf{u}^*$ defined by (9) may be larger than c and even if $u^* < c$ the transformation properties of the velocity $\mathbf{u}^*$ by Lorentz transformations will not in general be in accordance with the transformation equations (II. 55) for the velocity of a material particle. The energy density h may also be negative, thus corresponding to a negative mass density h/c^2.

Introducing in every system of coordinates four quantities S_ι, where

$$S_\iota = (\mathbf{S}, ich) = \frac{c}{i}\,T_{4\iota}, \tag{12}$$

we have, on account of (5),

$$\frac{\partial S_\iota}{\partial x_\iota} = 0 \tag{13}$$

in every system of coordinates The quantities S_ι do not, however, transform like the components of a four-vector. Let us now assume that $h > 0$ and $u^* < c$, i.e.

$$S_i S_i = S^2 - c^2h^2 < 0. \tag{14}$$

For every system of coordinates we may then define four quantities U_i^*, analogous to the four-velocity of a particle, by

$$U_i^* = \left(\frac{\mathbf{u}^*}{\sqrt{(1-u^{*2}/c^2)}}, \frac{ic}{\sqrt{(1-u^{*2}/c^2)}}\right). \tag{15}$$

These quantities will transform like the components of a time-like four-vector when and only when the velocity $\mathbf{u}^*$ transforms like a particle velocity. We shall now find the condition to be imposed on the tensor T_{ik} for this to be the case. Since

$$\sqrt{(1-u^{*2}/c^2)} = \sqrt{(1-S^2/h^2c^2)} = (-S_i S_i)^{\frac{1}{2}}/hc,$$

we have

$$\left.\begin{aligned} U_i^* &= \left(\frac{c\mathbf{S}}{(-S_l S_l)^{\frac{1}{2}}}, \frac{ihc^2}{(-S_l S_l)^{\frac{1}{2}}}\right) = \frac{cS_i}{(-S_l S_l)^{\frac{1}{2}}} \\ U_i^* U_i^* &= -c^2 \end{aligned}\right\}. \tag{16}$$

Now consider an infinitesimal Lorentz transformation

$$x_i' = x_i + \epsilon_{ik} x_k, \qquad \epsilon_{ik} = -\epsilon_{ki} \tag{17}$$

connecting the space-time coordinates of two systems of inertia S and S'. From the transformation laws for a tensor we then get

$$T_{4i}' = T_{4i} + \epsilon_{4l} T_{li} + \epsilon_{ik} T_{4k},$$

thus, on account of (12),

$$\left.\begin{aligned} S_i' &= S_i + \epsilon_{ik} S_k + \frac{c}{i} \epsilon_{4k} T_{ki} \\ S_i' S_i' &= S_i S_i + \frac{2c}{i} \epsilon_{4k} T_{ki} S_i \end{aligned}\right\}. \tag{17'}$$

Expanding in terms of the infinitesimal quantities ϵ_{ik} we get, neglecting terms of order higher than the first,

$$U_i^{*\prime} = \frac{cS_i'}{(-S_l' S_l')^{\frac{1}{2}}} = U_i^* + \epsilon_{ik} U_k^* + \frac{c^2 \epsilon_{4k}}{i(-S_l S_l)^{\frac{1}{2}}} \left(T_{ki} + \frac{U_i^* T_{kl} U_l^*}{c^2}\right). \tag{18}$$

Thus, in order that U_i^* shall transform like a four-vector the tensor T_{ik} must, for $i = 1, 2, 3$ and for all values of k, satisfy the condition

$$R_{ik} \equiv T_{ik} + \frac{T_{il} U_l^* U_k^*}{c^2} = 0. \tag{19}$$

For $i = 4$ the equation (19) is identically satisfied. This condition is, however, also sufficient, for, if (19) is satisfied in S, we have

$$U_i^{*\prime} = U_i^* + \epsilon_{ik} U_k^*, \tag{20}$$

R_{ik} will then transform like a tensor by the transformation from S to S' and (19) will therefore also be valid in S'. Since a finite Lorentz transformation may be composed of an infinite number of infinitesimal Lorentz transformations, (19) represents the general condition which T_{ik} must satisfy in order that the velocity of the energy $\mathbf{u}^*$ shall transform like a particle velocity

In general, the energy-momentum tensor of a physical system will not, of course, satisfy the condition (19) There are, however, a few cases in which this condition must be fulfilled for physical reasons. In the trivial case of a system consisting of incoherent matter without any external forces, for instance, we have $T_{ik} = \theta_{ik} = \mu^0 U_i U_k$, i e.

$$S_i = \frac{c}{i}\mu^0 U_i U_4, \qquad U_i^* = \frac{cS_i}{(-S_l S_l)^{\frac{1}{2}}} = U_i,$$

and $$R_{ik} = T_{ik} + \frac{T_{il} U_l^* U_k^*}{c^2} = \mu^0 U_i U_k + \frac{\mu^0 U_i U_l U_l U_k}{c^2} = 0.$$

In this case the propagation velocity of the energy is, of course, identical with the velocity of the matter For an elastic body, however, the condition (19) will in general not be fulfilled, as we shall see in § 65. In a later section (§ 76) we shall meet another important case in which the condition (19) must be satisfied for physical reasons.

63. Four-momentum and angular momentum four-tensor for a closed system

Putting $$g_i = \frac{1}{ic} T_{i4} = \left(\mathbf{g}, \frac{i}{c} h\right), \tag{21}$$

the symmetry condition (7) may be written

$$g_i = \frac{1}{c^2} S_i, \tag{22}$$

and, on account of (13), we have also

$$\frac{\partial g_i}{\partial x_i} = 0. \tag{23}$$

It should be remembered that the quantities g_i and S_i are not four-vectors.

Let us now assume that the system considered is finite so that all components T_{ik} are zero outside a certain region in physical space. If we multiply (1) by $dx_1 dx_2 dx_3$ and integrate over the whole physical space for constant x_4, the first three terms in (1), which are partial derivatives of $T_{\iota\kappa}$ with respect to the space coordinates x_κ, will give zero. Hence we get

$$\frac{d}{dx_4}\int T_{\iota 4}\,dV = 0,$$

which shows that the four quantities

$$G_\iota = \int g_\iota\,dV = \left(\mathbf{G}, \frac{i}{c}H\right) \tag{24}$$

are constant in time **G** and H represent the total linear momentum and the total energy of the system, respectively. Another consequence of (1) is that the quantities G_ι transform like the components of a four-vector. the four-momentum vector. This is seen in the following way.

Let a_ι be an arbitrary but constant four-vector. The four-vector

$$b_k = a_\iota T_{\iota k} \tag{25}$$

will then satisfy the equation

$$\frac{\partial b_k}{\partial x_k} = 0 \tag{26}$$

on account of (1).

If we multiply (26) by $d\Sigma = \frac{1}{i}\,dx_1 dx_2 dx_3 dx_4$ and integrate over a finite region Σ in (3+1)-space, we get by means of the generalized Gauss theorem (IV. 191)

$$0 = \int \frac{\partial b_k}{\partial x_k}\,d\Sigma = \int\limits_{\Omega} b_k\,dV_k, \tag{27}$$

where Ω is the three-dimensional boundary of the region Σ. Since we have to deal with a finite system, the region in (3+1)-space where $T_{\iota k}$ is different from zero, i e. the time track of the system, will have the form of a tube with a finite cross-section in the space-like directions

Consider two arbitrary coordinate systems S and S' in (3+1)-space and two hyperplanes Ω_1 and Ω_2 defined by the conditions

$$x_4 = \text{constant}$$

and

$$x_4' = \text{constant},$$

respectively, the values of the constants being arbitrary. For the region Σ we now choose a domain bounded by the hyperplanes Ω_1, Ω_2 and by a cylindrical surface Ω_3 enclosing the tube in which $T_{\iota k} \neq 0$ (Fig. 15) The three-dimensional hypersurface Ω is thus composed of the parts

Ω_1, Ω_2, and Ω_3. The contribution to the integral $\int\limits_{\Omega} b_k\, dV_k$ from the cylinder Ω_3 is zero, since T_{ik} and b_k are zero on Ω_3. Thus, we get

$$\int\limits_{\Omega_1} b_k\, dV_k + \int\limits_{\Omega_2} b_k\, dV_k = 0 \tag{28}$$

The two integrals in (28) are invariant and may thus be calculated in any system of coordinates We choose to calculate the first integral in S and

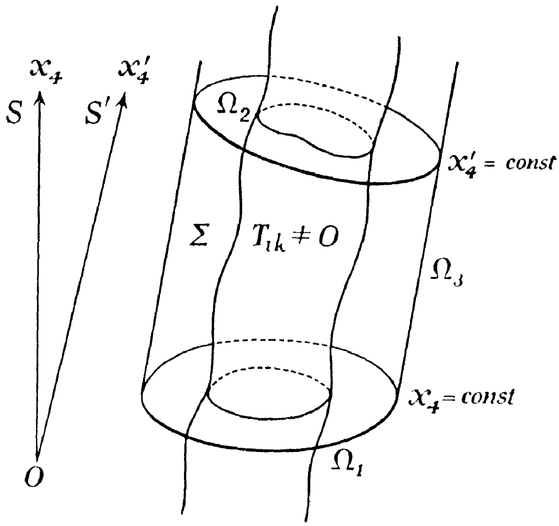

FIG 15.

the second in S' If we suppose the events on Ω_2 to be later in time than the events on Ω_1, the outward normal on Ω_1 is pointing in the direction of the negative time axis and, according to (IV. 191′), the components of dV_k are

$$dV_k = (0, 0, 0, +i\, dx_1\, dx_2\, dx_3)$$

Thus we get

$$\int\limits_{\Omega_1} b_k\, dV_k = +i \int b_4(x_1, x_2, x_3, x_4)\, dx_1\, dx_2\, dx_3, \tag{29}$$

where x_4 is kept constant in this integration

In the second integral we have similarly

$$dV'_k = (0, 0, 0, -i\, dx'_1\, dx'_2\, dx'_3),$$

hence

$$\int\limits_{\Omega_2} b_k\, dV_k = \int\limits_{\Omega_2} b'_k\, dV'_k = -i \int b'_4(x'_1, x'_2, x'_3, x'_4)\, dx'_1\, dx'_2\, dx'_3. \tag{30}$$

Using (29) and (30) in (28) we get

$$\int b_4\, dV = \int b'_4\, dV',$$

or, by means of (25), (21), and (24),

$$a_\iota G_\iota = a'_\iota G'_\iota = \text{invariant.} \tag{31}$$

Since this equation must hold for an arbitrary constant vector a_ι it follows that the connexion between G_ι and G'_ι is given by the transformation equation for a vector:

$$G'_\iota = \alpha_{\iota k} G_k. \tag{31'}$$

For the validity of this proof of the vector character of G_ι it is essential that $T_{\iota k}$ is everywhere regular. If $T_{\iota k}$ has a singularity along a certain world line, we would have to exclude this line from the region Σ by a certain surface Ω_4 which then would give a contribution to the integral $\int\limits_{\Omega} b_k\, dV_k$ in (27).

Since G_ι is a vector, $G_\iota G_\iota$ is an invariant, and we may now define the total proper mass M_0 of the system by the equation

$$G_\iota G_\iota = -M_0^2 c^2, \tag{32}$$

by analogy with (IV 51), holding for a material particle.

By means of (1) and (6) we get further

$$\frac{\partial}{\partial x_l}(x_\iota T_{kl} - x_k T_{\iota l}) = \delta_{\iota l} T_{kl} - \delta_{kl} T_{\iota l} = T_{k\iota} - T_{\iota k} = 0. \tag{33}$$

Integrating this equation over the whole physical space we find by arguments similar to those used before

$$\frac{d}{dx_4}\int (x_\iota T_{k4} - x_k T_{\iota 4})\, dV = 0,$$

which shows that the six quantities

$$M_{\iota k} = \int (x_\iota g_k - x_k g_\iota)\, dV = -M_{k\iota} \tag{34}$$

are constant in time. This result depends essentially on the symmetry of the energy-momentum tensor. By a method similar to that used in the proof of the vector character of G_ι it now follows from (1) that the quantities $M_{\iota k}$ transform like the components of an antisymmetrical tensor: the angular momentum four-tensor with respect to the arbitrary origin of the coordinate system.

The spatial part of this tensor $M_{\iota\kappa}$ is, according to (IV. 98), dual to an axial vector

$$\mathbf{M} = \int (\mathbf{x} \times \mathfrak{g})\, dV, \tag{35}$$

which is equal to the constant total angular momentum vector of the closed system.

64. Centre of mass†

We may assume that G_i for any physical system is a time-like vector so that M_0 defined by (32) is a real quantity (cf § 30) In this case it is always possible to find a system of inertia S^0, the 'rest system', in which the total linear momentum $\mathbf{G}^0 = 0$, so that on account of (32) we have for the components of G_i in S^0,

$$G_i^0 = (0, 0, 0, iM_0 c) \tag{36}$$

Exactly as in the case of a material particle the velocity $\mathbf{u}$ of the rest system S^0 relative to S is then

$$\mathbf{u} = c^2\mathbf{G}/H \tag{37}$$

In Newtonian mechanics the centre of mass of a physical system with the mass density $\mu = \mu(\mathbf{x}, t)$ is defined as a point with the coordinate vector

$$\mathbf{X} = \frac{1}{M}\int \mu(\mathbf{x}, t)\ \mathbf{x}\, dV, \tag{38}$$

where $M = \int \mu\, dV$ is the total mass of the system In relativistic mechanics the mass density is connected with the energy density by the equation (11) We may then define the centre of mass by the equations (38) and (11) As we shall see in a moment, the point defined in this way will, however, in general depend on the system of coordinates which is used in the evaluation of the integrals in (38), i e each system of inertia S has its own centre of mass $C(S)$ which depends on the system S In S its coordinate vector $\mathbf{X}(C(S)) = \mathbf{X}(S)$ is, according to (38), (11), and (24), defined by

$$\mathbf{X}(S) = \frac{1}{H}\int h(\mathbf{x}, t)\ \mathbf{x}\, dV \tag{39}$$

From (23) we now get in every system of coordinates

$$\frac{\partial(g_i x_k)}{\partial x_i} = g_i\,\delta_{ik} = g_k,$$

and by integration over the whole 3-space

$$\frac{d}{dx_4}\int g_4 x_k\, dV = G_k$$

Since H is constant in time this equation for $k = 1, 2, 3$ shows that the centre of mass defined by (39) is moving with the constant velocity $c^2\mathbf{G}/H$ relative to S, i e. with the same velocity $\mathbf{u}$ as the rest system S^0. Thus all the different mass centres $C(S)$ are at rest in the system S^0

† A D Fokker, *Relativiteitstheorie*, Groningen, 1929, p 170, A Papapetrou, *Praktika Acad d Athenes*, **14**, 540 (1939), C Møller, *Comm Dublin Inst for Advanced Studies*, Ser A, No 5 (1949), O C de Beauregard, *Relativité restreinte*, chap IX, Paris, 1949, M H L Pryce, *Proc Roy Soc* A, **195**, 62 (1948)

One of the centres of mass plays a distinguished role, viz. the point $C^0 = C(S^0)$ which is the centre of mass in the rest system itself, it may be called the proper centre of mass. If $X_i = (\mathbf{X}, X_4)$ are the space-time coordinates of the proper centre of mass C^0 in an arbitrary system of coordinates, the $X_i = X_i(\tau)$ will be linear functions of the proper time τ of this point. Further, if

$$U_i = \frac{dX_i}{d\tau} = \left(\frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}}, \frac{ic}{\sqrt{(1-u^2/c^2)}}\right)$$

denotes the four-velocity of C^0, we have, on account of (37),

$$G_i = M_0 U_i. \tag{40}$$

The dependence of G_i on the velocity of the proper centre of mass is thus the same as for a material particle.

Now denoting the relative angular momentum four-tensor m_{ik} with respect to the proper centre of mass C^0 by

$$m_{ik} = \int [(x_i - X_i)g_k - (x_k - X_k)g_i]\, dV = M_{ik} - (X_i G_k - X_k G_i), \tag{41}$$

we get, by differentiation with respect to the variables x_4 and τ,

$$\frac{dm_{ik}}{dx_4} = 0, \qquad \frac{dm_{ik}}{d\tau} = -U_i G_k + U_k G_i = 0$$

on account of (40). The two space vectors $\mathbf{m}$ and $\mathbf{n}$ defined by

$$\left.\begin{aligned} \mathbf{m} &= (m_{23}, m_{31}, m_{12}) \\ -i\mathbf{n} &= (m_{14}, m_{24}, m_{34}) \end{aligned}\right\} \tag{42}$$

are thus constants of the motion. We may therefore in (41) choose $x_4 = X_4$ and, by means of (42) and (41), we get

$$\left.\begin{aligned} \mathbf{m} &= \int ((\mathbf{x}-\mathbf{X})\times\mathbf{g})\, dV\Big|_{x_4 = X_4} \\ -\frac{\mathbf{n}}{c} &= \int \frac{h}{c^2}(\mathbf{x}-\mathbf{X})\, dV\Big|_{x_4 = X_4} = \int \mu(\mathbf{x}-\mathbf{X})\, dV\Big|_{x_4 = X_4} \end{aligned}\right\}. \tag{43}$$

Thus $\mathbf{m}$ is the relative angular momentum vector with respect to the proper centre of mass, the inner angular momentum, while $-\mathbf{n}/c$ is the moment of mass with respect to the same point. Solving the second equation (43) with respect to $\mathbf{X}$ we get

$$\mathbf{X} = \frac{1}{H}\int h\mathbf{x}\, dV\Big|_{x_4 = X_4} + c\mathbf{n}/H = \mathbf{X}(S) + \frac{c\mathbf{n}}{H}, \tag{44}$$

where $\mathbf{X}(S)$ and $\mathbf{X}(S^0) = \mathbf{X}$ are simultaneous coordinate vectors of the centre of mass $C(S)$ in S and the proper centre of mass $C(S^0)$.

From (44) we see that the different centres will coincide only if $\mathbf{n}$, and therefore m_{ik}, is zero in every system of inertia, i e when the system considered has no inner angular momentum If (44) is written in the rest system S^0 we get

$$\mathbf{n}^0 = 0 \quad \text{or} \quad m^0_{i4} = 0, \tag{45}$$

since by definition $\mathbf{X}^0 = \mathbf{X}^0(S^0)$. This condition is equivalent to the covariant equation

$$m_{ik}G_k = 0, \tag{46}$$

as is seen from (36) when equation (46) is written in the system S^0. Thus this equation expresses in a covariant way that the proper centre of mass is the centre of mass in its own rest system

When we choose the same orientation of the spatial axes in S as in S^0 we get, by means of the transformation equations (IV. 81′) for an antisymmetrical tensor, on account of (45),

$$\mathbf{n} = \frac{\mathbf{v}\times\mathbf{m}^0}{\sqrt{(c^2-v^2)}}, \tag{47}$$

where $\mathbf{v} = -\mathbf{u}$ is the velocity of S relative to S^0 and $\mathbf{m}^0$ is the inner angular momentum vector in the rest system

The difference between simultaneous positions of the centre of mass in S and the proper centre of mass is, according to (44) and (47), given by the time-independent space vector

$$\mathbf{a}(S) = \mathbf{X}(S)-\mathbf{X} = -c\mathbf{n}/H = (\mathbf{m}^0\times\mathbf{v})/M_0c^2, \tag{48}$$

where we have used the relations

$$H = \frac{M_0c^2}{\sqrt{(1-u^2/c^2)}} = \frac{M_0c^2}{\sqrt{(1-v^2/c^2)}}$$

following from (40).

Since the transformation from S to S^0 is given by a Lorentz transformation without rotation, and since $\mathbf{a}$ is perpendicular to the relative velocity $\mathbf{v}$, the distance between the two centres, mentioned above, in the rest system S^0 is also given by (48).

In the rest system S^0 all mass centres $C(S)$ obtained by varying S or $\mathbf{v}$ in (48) form a two-dimensional circular disk perpendicular to the angular momentum vector $\mathbf{m}^0$ with centre at the proper centre of mass C^0 and with radius

$$\rho = \frac{|\mathbf{m}^0|}{M_0c}. \tag{49}$$

In the non-relativistic limit $c\to\infty$ the radius of the disk tends to zero and we are left with one mass centre only, the Newtonian centre of gravity, but in the relativistic case we have in general not one centre of mass but the disk of mass centres mentioned above, the centre of which is the proper centre of mass. Only if the system has no inner angular momentum is the radius (49) of the disk zero. It is true that the radius (49) for all macroscopic systems is very small compared with the dimensions of the systems. For the earth, for instance, we have

$$\rho_{\text{earth}} = \frac{|\mathbf{m}^0|}{M_0 c} \approx 10 \text{ metres.} \tag{50}$$

For systems of atomic dimensions, however, the radius of the disk of mass centres may be comparable with the dimensions of the system.

From the above considerations we can draw a certain conclusion regarding the dimensions of a system with given inner angular momentum $\mathbf{m}^0$ and proper mass M_0. Consider an arbitrary physical system which in the rest system S^0 lies entirely inside a sphere with centre at the proper centre of mass C^0 and radius r, i.e. a system for which all components of the energy-momentum tensor are zero outside this sphere. If we further assume that *the energy density h is positive everywhere in all systems of inertia*, it is clear that the whole of the disk of mass centres must lie inside the sphere; for if we consider an arbitrary point, say $C(S)$, on the disk, this point will in the system of coordinates S be a centre of mass, and since h is positive it must then lie inside the physical system. We thus get

$$r \geqslant \frac{|\mathbf{m}^0|}{M_0 c} \tag{51}$$

Thus, *a system with positive energy density and with a given inner angular momentum m^0 and a given rest mass M_0 must always have a finite extension in accordance with* (51). If the system is smaller, h cannot be everywhere positive in all systems of inertia.

65. The fundamental equations of mechanics in elastic continua

In Chapter IV, §§ 50–51, we have treated the mechanics of incoherent matter under the influence of given external forces. We shall now consider the case of an elastic body with no external forces. The sole forces acting in the body are then the elastic forces between neighbouring parts of the matter due to the deformation of the matter. We have thus to deal with a closed system which is a special case of the general systems considered in § 62, and the equations (1)–(11) must be valid for the total energy-momentum tensor T_{ik} of this mechanical system. The mechanical energy-momentum tensor has, however, especially simple properties which we shall now establish.

Consider an infinitesimal face element $d\sigma$ with a directed normal defined by a unit vector $\mathbf{n}$, at a definite point p in space. The matter on either side of this face element experiences a force which is proportional to $d\sigma$. The force acting on the side to which the normal points will be called $\mathbf{t}(\mathbf{n})\,d\sigma$, and, since action and reaction are equal, the force on the other side $\mathbf{t}(-\mathbf{n})\,d\sigma$ must then be $-\mathbf{t}(\mathbf{n})\,d\sigma$ If $\mathbf{n}^{(1)}$, $\mathbf{n}^{(2)}$, $\mathbf{n}^{(3)}$ are unit vectors in the directions of the Cartesian axes, we have

$$\mathbf{t}(\mathbf{n}) = \mathbf{t}(\mathbf{n}^{(1)})n_1+\mathbf{t}(\mathbf{n}^{(2)})n_2+\mathbf{t}(\mathbf{n}^{(3)})n_3, \tag{52}$$

where n_1, n_2, n_3 are the components of the unit vector $\mathbf{n}$ The equation (52) is obtained by a consideration of the infinitesimal piece of matter

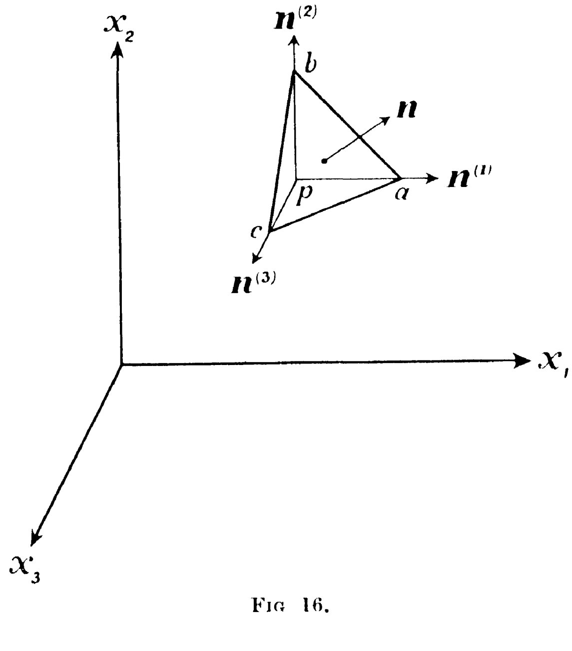

FIG 16.

contained in the pyramid $abcp$ of Fig 16 If $d\sigma$ is the area of the triangle abc, the area of the triangles pbc, pca, and pab are $n_1\,d\sigma$, $n_2\,d\sigma$, $n_3\,d\sigma$, respectively, and the total elastic force on this piece of matter will be

$$-\mathbf{t}(\mathbf{n})\,d\sigma+\mathbf{t}(\mathbf{n}^{(1)})n_1\,d\sigma+\mathbf{t}(\mathbf{n}^{(2)})n_2\,d\sigma+\mathbf{t}(\mathbf{n}^{(3)})n_3\,d\sigma$$

This force must be equal to the change of momentum of the matter per unit time, i.e. $d(\mathbf{g}\,\delta V)/dt$, where $\mathbf{g}$ is the momentum density and δV the volume of the pyramid In the limit of an infinitely small volume, δV tends to zero faster than $d\sigma$, therefore

$$\frac{1}{d\sigma}\frac{d}{dt}(\mathbf{g}\,\delta V) \to 0,$$

which leads immediately to the equation (52) If the components of the vectors $\mathbf{t}(\mathbf{n}^{(\kappa)})$ are denoted by $t_{\iota\kappa}$, (52) may be written

$$\mathbf{t}_\iota(\mathbf{n}) = t_{\iota\kappa} n_\kappa \tag{53}$$

Since $t_\iota(\mathbf{n})$ and n_κ are the components of space vectors, the quantities $t_{\iota\kappa}$ must transform like the components of a space tensor by rotations of the Cartesian axes. $t_{\iota\kappa}$ is the elastic stress tensor, sometimes called the relative stress tensor, in contrast to the space part $T_{\iota\kappa}$ of the total energy-momentum tensor T_{ik} which is called the absolute stress tensor †

The total elastic force $\mathbf{F}$ acting on the matter inside a closed surface σ is now equal to

$$-\int_\sigma \mathbf{t}(\mathbf{n})\, d\sigma,$$

where $\mathbf{n}$ is the outward normal to the surface element $d\sigma$.

The components F_ι of this force may, by means of (53) and Gauss's theorem (IV. 192), be written

$$F_\iota = -\int_\sigma t_{\iota\kappa} n_\kappa \, d\sigma = -\int_\Omega \frac{\partial t_{\iota\kappa}}{\partial x_\kappa} dV, \tag{54}$$

where the integration on the right-hand side is extended over the interior of the closed surface σ Thus we can define an elastic force density $\mathbf{f}$ such that

$$F_\iota = \int_\Omega f_\iota \, dV, \tag{55}$$

and a comparison of (54) and (55) shows that the elastic force density and the relative stress tensor are connected by the equation

$$f_\iota = -\frac{\partial t_{\iota\kappa}}{\partial x_\kappa}. \tag{56}$$

The motion of an infinitesimal piece of matter with the volume δV is now determined by the equations of motion

$$\frac{d}{dt}(g_\iota \, \delta V) = f_\iota \, \delta V = -\frac{\partial t_{\iota\kappa}}{\partial x_\kappa} \delta V, \tag{57}$$

where $\mathbf{g}$ is the momentum density and d/dt denotes the substantial time derivative By means of (IV 201) and (IV. 204) we get

$$\frac{d}{dt}(g_\iota \, \delta V) = \frac{dg_\iota}{dt} \delta V + g_\iota \frac{d}{dt} \delta V = \left(\frac{\partial g_\iota}{\partial t} + \frac{\partial g_\iota}{\partial x_\kappa} u_\kappa\right) \delta V + g_\iota \, \delta V \cdot \frac{\partial u_\kappa}{\partial x_\kappa}$$
$$= \left(\frac{\partial g_\iota}{\partial t} + \frac{\partial (g_\iota u_\kappa)}{\partial x_\kappa}\right) \delta V, \tag{58}$$

† M von Laue, *Die Relativitatstheorie* (3rd ed , Braunschweig 1919), vol 1, § 29; *Ann d Phys* **35**, 524 (1911).

where the u_κ are the components of the velocity $\mathbf{u}$ of the matter at the place considered. From (57) and (58) we then get

$$\frac{\partial g_\iota}{\partial t}+\frac{\partial}{\partial x_\kappa}(g_\iota u_\kappa+t_{\iota\kappa}) = 0. \tag{59}$$

On the other hand, the law of conservation of momentum is also expressed by (4); thus we obtain the following connexion between the absolute and the relative stress tensors

$$T_{\iota\kappa} = t_{\iota\kappa}+g_\iota u_\kappa. \tag{60}$$

In order to find an explicit expression for the momentum density we shall use the connexion (8) between $\mathfrak{g}$ and the energy flux $\mathbf{S}$:

$$\mathfrak{g} = \mathbf{S}/c^2. \tag{61}$$

The total work done by elastic forces on the matter inside a closed surface σ per unit time is

$$A = -\int\limits_\sigma (\mathbf{t}(\mathbf{n}).\mathbf{u})\,d\sigma = -\int\limits_\sigma t_{\iota\kappa}n_\kappa u_\iota\,d\sigma = -\int\limits_\Omega \frac{\partial(u_\iota t_{\iota\kappa})}{\partial x_\kappa}\,dV,$$

where the integration in the last integral is extended over the interior Ω of the surface σ. The work done on an infinitesimal piece of matter of volume δV is thus

$$\delta A = -\frac{\partial(u_\iota t_{\iota\kappa})}{\partial x_\kappa}\,\delta V. \tag{62}$$

This must be equal to the increase per unit time of the energy inside δV which is

$$\frac{d}{dt}(h\,\delta V) = \left(\frac{\partial h}{\partial t}+\frac{\partial h}{\partial x_\kappa}u_\kappa\right)\delta V+h\,\delta V\frac{\partial u_\kappa}{\partial x_\kappa} = \left[\frac{\partial h}{\partial t}+\frac{\partial}{\partial x_\kappa}(hu_\kappa)\right]\delta V, \tag{63}$$

h being the total energy density including the elastic energy. Thus we get from (63) and (62)

$$\frac{\partial h}{\partial t}+\frac{\partial}{\partial x_\kappa}(hu_\kappa+u_\iota t_{\iota\kappa}) = 0. \tag{64}$$

A comparison of (5) and (64) shows that the total energy flux is given by

$$\mathbf{S} = h\mathbf{u}+(\mathbf{u}.\mathbf{t}), \tag{65}$$

where $(\mathbf{u}\ \mathbf{t})$ is a space vector with components $(\mathbf{u}.\mathbf{t})_\kappa = u_\iota t_{\iota\kappa}$ Thus, besides the convection current $h\mathbf{u}$ there is an extra transport of energy due to the work done by the elastic forces From (61) and (11) we then get for the total momentum density

$$\mathfrak{g} = \frac{\mathbf{S}}{c^2} = \mu\mathbf{u}+\frac{(\mathbf{u}.\mathbf{t})}{c^2}, \tag{66}$$

where $\mu = h/c^2$ is the total mass density including the mass of the elastic energy On account of the last term in (66) the momentum density vector has not in general the same direction as the direction of motion of the matter, hence

$$g_\iota u_\kappa \neq g_\kappa u_\iota$$

Since the law of conservation of angular momentum requires $T_{\iota\kappa} = T_{\kappa\iota}$ (see § 63), we get from (60)

$$t_{\iota\kappa} - t_{\kappa\iota} = -g_\iota u_\kappa + g_\kappa u_\iota = (-(\mathbf{u}.\mathbf{t})_\iota u_\kappa + (\mathbf{u}.\mathbf{t})_\kappa u_\iota)/c^2 \neq 0,$$

i e the relative stress tensor is not symmetrical

Only in the momentary rest system S^0 of the matter at the point considered, we have $\mathbf{u}^0 = 0$ and thus, on account of (60), (65), and (66),

$$t^0_{\iota\kappa} = T^0_{\iota\kappa} = T^0_{\kappa\iota} = t^0_{\kappa\iota}, \qquad S^0_\iota = \mathbf{g}^0_\iota = T^0_{\iota 4} = 0, \qquad T^0_{44} = -h^0, \quad (67)$$

where h^0 is the rest energy density.

The mechanical energy-momentum tensor satisfies the equation

$$T_{ik} U_k = -h^0 U_i, \quad (68)$$

where U_i is the four-velocity of the matter The validity of the covariant equation (68) follows at once from (67) if it is written down in the rest system, where $U^0_i = (0, 0, 0, ic)$. The relation (68) is characteristic of a pure mechanical energy-momentum tensor and contains equation (IV. 246) as a special case. If we multiply (68) by U_i we get the following expression for the invariant rest energy density

$$h^0 = U_i T_{ik} U_k/c^2. \quad (69)$$

For $i = \iota = 1, 2, 3$, (68) gives

$$T_{\iota\kappa} U_\kappa + T_{\iota 4} U_4 = -h^0 U_\iota,$$

or, by means of (60), (3), and (IV 39),

$$(t_{\iota\kappa} + g_\iota u_\kappa) u_\kappa - c^2 g_\iota = -h^0 u_\iota.$$

Solving with respect to g_ι we get a new expression for the momentum density

$$\mathbf{g} = \frac{\mu^0 \mathbf{u} + (1/c^2)(\mathbf{t}.\mathbf{u})}{1 - u^2/c^2}, \quad (70)$$

where $\mu^0 = h^0/c^2$ is the rest energy density and $(\mathbf{t}.\mathbf{u})$ is a space vector with components $t_{\iota\kappa} u_\kappa$. By means of (IV. 39), (70) may also be written

$$T_{\iota 4} = T_{4\iota} = icg_\iota = \mu^0 U_\iota U_4 + \frac{t_{\iota\lambda} U_\lambda U_4}{c^2}. \quad (71)$$

The equation (68) for $i = 4$ gives similarly, on account of (70),

$$h = h^0 + (\mathbf{g}.\mathbf{u}) = \frac{h^0 + (\mathbf{u}.\mathbf{t}.\mathbf{u})/c^2}{1 - u^2/c^2}, \quad (72)$$

with $(\mathbf{u}.\mathbf{t}.\mathbf{u}) = u_\iota t_{\iota\kappa} u_\kappa$.

(72) may also be written

$$T_{44} = -h = +\mu^0 U_4 U_4 - \frac{1}{c^2} U_\iota t_{\iota\kappa} U_\kappa. \tag{73}$$

The equation (69) gives similarly

$$h^0 = h\left(1-\frac{u^2}{c^2}\right) - \frac{1}{c^2}(\mathbf{u}.\mathbf{t}.\mathbf{u}), \tag{74}$$

which is identical with (72) The right-hand side of (74) is thus an invariant scalar

If we divide (74) by c^2 we get

$$\mu^0 = \mu\left(1-\frac{u^2}{c^2}\right) - \frac{1}{c^4}(\mathbf{u}\ \mathbf{t}.\mathbf{u}), \tag{75}$$

which is a generalization of the equation (IV 199) valid for incoherent matter

Comparing the two expressions (66) and (70) for $\mathfrak{g}$ we get, by means of (75), the following identity for the relative stress tensor $t_{\iota\kappa}$

$$(1-u^2/c^2)(\mathbf{u}.\mathbf{t}) = (\mathbf{t}.\mathbf{u}) - \mathbf{u}(\mathbf{u}\ \mathbf{t}.\mathbf{u})/c^2 \tag{76}$$

A closed system may always be divided into non-closed sub-systems in an infinite number of ways corresponding to a division of the energy-momentum tensor into separate parts. For instance, we may write

$$T_{ik} = \theta_{ik} + S_{ik}, \tag{77}$$

where

$$\theta_{ik} = \mu^0 U_i U_k \tag{78}$$

is the kinetic energy-momentum tensor By means of (60), (71), and (IV. 39) we get

$$T_{\iota\kappa} = t_{\iota\kappa} + g_\iota u_\kappa = \mu^0 U_\iota U_\kappa + t_{\iota\kappa} + \frac{t_{\iota\lambda} U_\lambda U_\kappa}{c^2}. \tag{79}$$

Thus we get for the components of the tensor S_{ik} from (79), (71), and (73)

$$\left.\begin{aligned} S_{\iota\kappa} = S_{\kappa\iota} &= T_{\iota\kappa} - \mu^0 U_\iota U_\kappa = t_{\iota\kappa} + \frac{t_{\iota\lambda} U_\lambda U_\kappa}{c^2} \\ S_{\iota 4} = S_{4\iota} &= T_{\iota 4} - \mu^0 U_\iota U_4 = \frac{t_{\iota\lambda} U_\lambda U_4}{c^2} \\ S_{44} &= T_{44} - \mu^0 U_4 U_4 = -\frac{1}{c^2} U_\iota t_{\iota\kappa} U_\kappa \end{aligned}\right\}. \tag{80}$$

On account of (68) the tensor S_{ik} satisfies the condition

$$S_{ik} U_k = (T_{ik} - \mu^0 U_i U_k) U_k = -h^0 U_i + \mu^0 c^2 U_i = 0, \tag{81}$$

a relation which also follows directly from the expression (80).

In the rest system we get from (80)

$$S^0_{\iota\kappa} = t^0_{\iota\kappa}, \qquad S^0_{\iota 4} = S^0_{4\iota} = 0 \tag{82}$$

Putting
$$f_i^{\text{elast}} = -\frac{\partial S_{ik}}{\partial x_k}, \tag{83}$$
the equations (1) may be written in the form
$$\frac{\partial \theta_{ik}}{\partial x_k} = f_i^{\text{elast}} \tag{84}$$
analogous to (IV 229) It should be noted that the force density $\mathbf{f}^{\text{elast}}$, defined by (83), is not in general identical with the elastic force $\mathbf{f}$ defined by (56) A simple calculation shows that $(f_i^{\text{elast}} U_i) \neq 0$ which, as shown in § 50, means that the proper mass of the system is not conserved This is also natural, since μ^0 includes the mass corresponding to the elastic energy which changes under the influence of the elastic forces While the tensor θ_{ik} satisfies the condition (19), the total mechanical energy-momentum tensor T_{ik} will in general not satisfy this condition

The relative stress tensor $t_{\iota\kappa}$ is connected with the internal deformation of the matter In the rest system S^0 this connexion is given by the equations of the usual theory of elasticity, thus for small deformations it is given by Hooke's law By means of the transformation properties of the stress tensor this connexion can be established in any system of inertia †

If the elastic body is subjected also to external non-mechanical forces described by a four-force density f_i^{ext} we have instead of (1)
$$\frac{\partial T_{ik}}{\partial x_k} = f_i^{\text{ext}}, \tag{85}$$
or by means of (77) and (83)
$$\frac{\partial \theta_{ik}}{\partial x_k} = f_i^{\text{elast}} + f_i^{\text{ext}} = f_i \tag{86}$$

66. Transformation of elastic stress, momentum density, and energy density

Let us assume the spatial axes in the coordinate systems S and S^0 to have the same orientation. Since the velocity of S^0 relative to S is the same as the velocity $\mathbf{u}$ of the matter at the point considered, the transformation coefficients α_{ik} are given by (IV 125), i e
$$\left.\begin{aligned} &\alpha_{\iota\kappa} = \delta_{\iota\kappa} + \frac{u_\iota u_\kappa}{u^2}(\gamma - 1) \\ &\alpha_{4\iota} = -\alpha_{\iota 4} = -\frac{i}{c}\gamma u_\iota, \qquad \alpha_{44} = \gamma \\ &\gamma = (1 - u^2/c^2)^{-\frac{1}{2}} \end{aligned}\right\}. \tag{87}$$

† G Herglotz, *Ann d Phys.* **36**, 493 (1911).

From the transformation equations of a tensor in connexion with (67) we get

$$T_{ik} = T^0_{lm}\alpha_{li}\alpha_{mk} = t^0_{\lambda\mu}\alpha_{\lambda i}\alpha_{\mu k} - h^0\alpha_{4i}\alpha_{4k}. \tag{88}$$

For $i = k = 4$ this gives

$$\left.\begin{aligned} h &= -t^0_{\lambda\mu}\alpha_{\lambda 4}\alpha_{\mu 4} + h^0\alpha^2_{44} = \gamma^2\frac{u_\lambda t^0_{\lambda\mu}u_\mu}{c^2} + \gamma^2 h^0 \\ h &= \frac{h^0 + (\mathbf{u}.\mathbf{t}^0.\mathbf{u})/c^2}{1 - u^2/c^2} = U_i T^0_{ik} U_k/c^2 \end{aligned}\right\}. \tag{89}$$

A comparison of (89) and (72) shows that we must have

$$(\mathbf{u}.\mathbf{t}.\mathbf{u}) = (\mathbf{u}.\mathbf{t}^0.\mathbf{u}). \tag{90}$$

For $i = \iota = 1, 2, 3$, $k = 4$, we get from (88), by means of (87),

$$g_\iota = \frac{T_{\iota 4}}{ic} = \frac{t^0_{\lambda\mu}\alpha_{\lambda\iota}\alpha_{\mu 4} - h^0\alpha_{4\iota}\alpha_{44}}{ic} = \frac{1}{c^2}\left(\delta_{\iota\lambda} + \frac{u_\iota u_\lambda}{u^2}(\gamma - 1)\right)\gamma u_\mu t^0_{\lambda\mu} + \frac{1}{c^2}\gamma^2 u_\iota h^0$$

$$= u_\iota\frac{\gamma^2}{c^2}\left[h^0 + \frac{(u_\lambda t^0_{\lambda\mu}u_\mu)}{u^2}\left(1 - \frac{1}{\gamma}\right)\right] + \frac{\gamma}{c^2}t^0_{\iota\mu}u_\mu, \tag{91}$$

or, in vector and tensor notation,

$$\mathbf{g} = \mathbf{u}\frac{\gamma^2}{c^2}\left[h^0 + \frac{(\mathbf{u}.\mathbf{t}^0.\mathbf{u})}{u^2}\left(1 - \frac{1}{\gamma}\right)\right] + \frac{\gamma}{c^2}(\mathbf{t}^0.\mathbf{u}). \tag{92}$$

In the same way we get from (88) with $i = \iota$, $k = \kappa$, and from (60) and (91)

$$t_{\iota\kappa} = t^0_{\iota\kappa} + u_\iota u_\lambda t^0_{\lambda\kappa}\frac{\gamma - 1}{u^2} - t^0_{\iota\lambda}u_\lambda u_\kappa\frac{\gamma - 1}{\gamma u^2} - u_\iota u_\kappa\frac{(u_\lambda t^0_{\lambda\mu}u_\mu)}{u^4}\frac{(\gamma - 1)^2}{\gamma}. \tag{93}$$

Introducing the notation $\mathbf{a} \circ \mathbf{b}$ for the direct product of the space vectors $\mathbf{a}$ and $\mathbf{b}$, which is a tensor with components $a_\iota b_\kappa$, the formula (93) may also be written in three-dimensional tensor notation

$$\mathbf{t} = \mathbf{t}^0 + \mathbf{u} \circ (\mathbf{u}.\mathbf{t}^0)\frac{(\gamma - 1)}{u^2} - (\mathbf{t}^0.\mathbf{u}) \circ \mathbf{u}\frac{\gamma - 1}{\gamma u^2} - \mathbf{u} \circ \mathbf{u}\frac{(\mathbf{u}.\mathbf{t}^0.\mathbf{u})}{u^4}\frac{(\gamma - 1)^2}{\gamma}. \tag{94}$$

By means of this transformation equation it is easily verified that the relation (90) is valid, further remembering that $t^0_{\iota\kappa}$ is symmetrical, i.e. $(\mathbf{t}^0.\mathbf{u}) = (\mathbf{u}.\mathbf{t}^0)$, a simple calculation shows that the expression (94) for $\mathbf{t}$ is in accordance with the equation (76).

In the special case where $\mathbf{u} = (u, 0, 0)$, i.e. where the motion of the matter at the point considered is parallel to the x-axis, the transformation equations (89), (92), and (93) reduce to

$$\left.\begin{aligned}
&h = \left(h^0 + \frac{u^2}{c^2} t^0_{xx}\right)\gamma^2 \\
&g_x = \gamma^2\left(\mu^0 + \frac{t^0_{xx}}{c^2}\right)u, \qquad g_y = \frac{\gamma t^0_{yx}}{c^2}u, \qquad g_z = \frac{\gamma t^0_{zx}}{c^2}u \\
&t_{xx} = t^0_{xx}, \qquad t_{xy} = \gamma t^0_{xy}, \qquad t_{xz} = \gamma t^0_{xz} \\
&t_{yx} = \frac{1}{\gamma} t^0_{yx}, \qquad t_{yy} = t^0_{yy}, \qquad t_{yz} = t^0_{yz} \\
&t_{zx} = \frac{t^0_{zx}}{\gamma}, \qquad t_{zy} = t^0_{zy}, \qquad t_{zz} = t^0_{zz}
\end{aligned}\right\}. \tag{95}$$

67. Perfect fluids

In a perfect fluid the force $\mathbf{t}(\mathbf{n})$ on a face element with normal $\mathbf{n}$ is parallel to $\mathbf{n}$, i.e.

$$\mathbf{t}(\mathbf{n}) = p\mathbf{n}, \tag{96}$$

where p is the normal pressure Thus we get from (53)

$$t_{\iota\kappa} = p\,\delta_{\iota\kappa}, \tag{97}$$

where p is the normal pressure.

In the rest system we have in particular

$$t^0_{\iota\kappa} = p^0\,\delta_{\iota\kappa}. \tag{98}$$

The transformation equations (93) in this case reduce to

$$p\,\delta_{\iota\kappa} = p^0\,\delta_{\iota\kappa}, \tag{99}$$

or

$$p = p^0, \tag{100}$$

i e the normal pressure is an invariant scalar †

From (80) and (97) we now get for a perfect fluid

$$\left.\begin{aligned}
&S_{\iota\kappa} = S_{\kappa\iota} = p\,\delta_{\iota\kappa} + \frac{p}{c^2}U_\iota U_\kappa \\
&S_{\iota 4} = S_{4\iota} = \frac{p}{c^2}U_\iota U_4 \\
&S_{44} = -\frac{p}{c^2}(U_\iota U_\iota) = \frac{p}{c^2}U_4 U_4 + p
\end{aligned}\right\}, \tag{101}$$

i e

$$S_{ik} = \frac{p}{c^2}U_i U_k + p\,\delta_{ik}, \tag{102}$$

where $p = p^0$ is the invariant pressure. This expression for S_{ik} also follows from the fact that (102) is identical with (82) in the rest system.

† M. Planck, *Berl. Ber*, p. 542 (1907), *Ann. d. Phys* **76**, 1 (1908).

From (102) it follows that the pressure p is one-third of the invariant diagonal sum of the tensor S_{ik}, i.e.

$$p = \tfrac{1}{3}S_{ii} \tag{103}$$

The total energy-momentum tensor for a perfect fluid is then, by (77) and (78),

$$T_{ik} = \left(\mu^0 + \frac{p}{c^2}\right) U_i U_k + p\,\delta_{ik}. \tag{104}$$

The pressure $p = p^0$ is a function of the rest density μ^0 and the temperature which is given by the equation of state of the fluid.

If the fluid is subjected to external forces with the four-force density f_i^{ext}, the equations of motion of the fluid may be written in the form (86) with

$$\left.\begin{aligned} f_i^{\text{elast}} &= -\frac{\partial S_{ik}}{\partial x_k} = -\frac{pU_i}{c^2}\frac{\partial U_k}{\partial x_k} - \frac{pU_k}{c^2}\frac{\partial U_i}{\partial x_k} - \frac{U_i U_k}{c^2}\frac{\partial p}{\partial x_k} - \frac{\partial p}{\partial x_i} \\ &= -\frac{pU_i}{c^2}\frac{\partial U_k}{\partial x_k} - \frac{p}{c^2}\frac{dU_i}{d\tau} - \frac{U_i}{c^2}\frac{dp}{d\tau} - \frac{\partial p}{\partial x_i} \end{aligned}\right\}. \tag{105}$$

Let us assume f_i^{ext} to be of the type (IV. 214) which satisfies the identity (IV. 216), $U_i f_i^{\text{ext}} = 0$. The action of these forces will then not give rise to any creation of proper mass. It is different, however, with the forces (105) for which

$$(U_i f_i^{\text{elast}}) = p\frac{\partial U_k}{\partial x_k} + \frac{dp}{d\tau} - U_i\frac{\partial p}{\partial x_i} = p\frac{\partial U_k}{\partial x_k} = p^0 \operatorname{div}^0 \mathbf{u}^0 \tag{106}$$

$\operatorname{div}^0 \mathbf{u}^0$ is the volume dilatation in the rest system, $-p^0 \operatorname{div}^0 \mathbf{u}^0$ represents therefore the increase in elastic potential energy density per unit time in the rest system. Since μ^0 includes the mass corresponding to elastic potential energy, $-\dfrac{p}{c^2}\dfrac{\partial U_k}{\partial x_k}$ thus represents the rate of creation of proper mass density in the rest system. Multiplying (86) by U_i and summing we now get the equation analogous to (IV. 226)

$$\frac{\partial \mu^0 U_k}{\partial x_k} = -\frac{U_i f_i^{\text{elast}}}{c^2} = -p\frac{(\partial U_k/\partial x_k)}{c^2}, \tag{107}$$

which just expresses that
$$-\frac{p}{c^2}\frac{\partial U_k}{\partial x_k} \tag{108}$$

represents the source density of proper mass. Only if the fluid may be considered as incompressible do we have $\partial U_k/\partial x_k = \operatorname{div}^0 \mathbf{u}^0 = 0$, and the proper mass will be conserved.

By means of (107) we now get

$$\frac{\partial \theta_{ik}}{\partial x_k} = \mu^0 U_k \frac{\partial U_i}{\partial x_k} + U_i \frac{\partial(\mu^0 U_k)}{\partial x_k} = \mu^0 \frac{dU_i}{d\tau} - \frac{pU_i}{c^2}\frac{\partial U_k}{\partial x_k}. \tag{109}$$

Thus, using (109) and (105) in (86), we get the following equations of motion for a perfect fluid.

$$\left(\mu^0 + \frac{p}{c^2}\right)\frac{dU_i}{d\tau} = f_i^{\text{ext}} - \frac{\partial p}{\partial x_i} - \frac{U_i}{c^2}\frac{dp}{d\tau}. \tag{110}$$

For the energy density and momentum density we get from (66), (70), (89), and (97), or directly from (104),

$$\mathbf{g} = \left(\mu + \frac{p}{c^2}\right)\mathbf{u} = \frac{(\mu^0 + p^0/c^2)\mathbf{u}}{1-u^2/c^2}, \qquad h = \frac{h^0 + p^0 u^2/c^2}{1-u^2/c^2}, \qquad p = p^0. \tag{111}$$

If μ^0, p^0, and $\mathbf{u}$ are constant throughout the elastic body, we get by integration of (111) over the whole volume $V = V^0\sqrt{(1-u^2/c^2)}$ of the body

$$\left.\begin{aligned} \mathbf{G} = \mathbf{g}V &= \frac{\dfrac{h^0+p^0}{c^2}V\mathbf{u}}{1-\dfrac{u^2}{c^2}} = \frac{\dfrac{(h^0+p^0)V^0}{c^2}\mathbf{u}}{\sqrt{\left(1-\dfrac{u^2}{c^2}\right)}} = \frac{\dfrac{H^0+p^0V^0}{c^2}\mathbf{u}}{\sqrt{\left(1-\dfrac{u^2}{c^2}\right)}} \\ H = hV &= \frac{H^0 + \dfrac{p^0V^0}{c^2}u^2}{\sqrt{\left(1-\dfrac{u^2}{c^2}\right)}} \end{aligned}\right\}. \tag{112}$$

From these equations we see that the total momentum and energy, i.e. the quantities $(\mathbf{G}, (i/c)H)$, do not form a four-vector in this case. This is not in contradiction with our general result in § 63, because the system is not a closed system. In order that the quantities μ^0, p^0, $\mathbf{u}$ can be constant throughout the body, the fluid must be contained in a vessel, the walls of which will act on the system with forces which are not included in the energy-momentum tensor (104). (See Chapter VII.) From (112) we find, however,

$$\mathbf{G} = \frac{H+pV}{c^2}\mathbf{u}, \qquad H+pV = \frac{H^0+p^0V^0}{\sqrt{(1-u^2/c^2)}}, \tag{113}$$

which shows that the system has the same momentum as a particle with energy $E = H+pV$ and rest energy $E^0 = H^0+p^0V^0$. The quantities

$$p_i = \left(\mathbf{G}, \frac{i}{c}(H+pV)\right) \tag{114}$$

thus transform like the components of a four-vector.

68. Scalar meson fields. General field theory

While the forces between the atomic nuclei and the outer electrons are properly described by electromagnetic fields, the characteristic short-range property of the forces between the constituent particles of the nuclei indicates that the nuclear forces are of an essentially non-electromagnetic nature. In order to account for the nuclear forces Yukawa† introduced the so-called meson fields. The simplest type of meson field is the scalar field described by an invariant scalar field function $\Psi(x_i)$ satisfying the field equation

$$\frac{\partial^2\Psi}{\partial x_i\,\partial x_i}-\kappa^2\Psi=0 \quad \text{or} \quad \Box\Psi-\kappa^2\Psi=0 \tag{115}$$

Here κ is a constant connected with the range of the nuclear forces and in the case of a 'neutral' meson field Ψ is a real function of the space-time coordinates (x_i).

Introducing the notation $\Psi_i=\dfrac{\partial\Psi}{\partial x_i}$,

the field equations can also be written

$$\frac{\partial\Psi_i}{\partial x_i}-\kappa^2\Psi=0 \tag{116}$$

These equations may be derived from a variational principle

$$\delta\int\mathfrak{L}(\Psi,\Psi_i)\,d\Sigma=0, \tag{117}$$

where
$$\mathfrak{L}=-\tfrac{1}{2}(\Psi_i\Psi_i+\kappa^2\Psi^2). \tag{118}$$

In fact, if the variation $\delta\Psi=\delta\Psi(x_i)$ of Ψ is assumed to vanish at the boundary of the arbitrary four-dimensional region of integration we have, since $\delta\Psi_i=(\partial/\partial x_i)\delta\Psi$,

$$\left.\begin{aligned}\delta\int\mathfrak{L}\,d\Sigma=\int\delta\mathfrak{L}\,d\Sigma&=\int\left(\frac{\partial\mathfrak{L}}{\partial\Psi}\delta\Psi+\frac{\partial\mathfrak{L}}{\partial\Psi_i}\delta\Psi_i\right)d\Sigma\\&=\int\left[\frac{\partial\mathfrak{L}}{\partial\Psi}-\frac{\partial}{\partial x_i}\left(\frac{\partial\mathfrak{L}}{\partial\Psi_i}\right)\right]\delta\Psi\,d\Sigma\end{aligned}\right\}. \tag{119}$$

Now, as this expression is to be zero for any variation of Ψ of the type considered, we get

$$\frac{\partial\mathfrak{L}}{\partial\Psi}-\frac{\partial}{\partial x_i}\left(\frac{\partial\mathfrak{L}}{\partial\Psi_i}\right)=0, \tag{120}$$

which are the Euler equations corresponding to the variational principle (117). With the expression (118) for $\mathfrak{L}$ (120) is identical with the field equations (116).

† H. Yukawa, *Proc. Math.-Phys. Soc. Japan*, **17**, 48 (1935).

The energy-momentum tensor of the scalar meson field is given by

$$T_{ik} = \Psi_i \Psi_k + \mathfrak{L}\,\delta_{ik} = \Psi_i \Psi_k - \tfrac{1}{2}(\Psi_l \Psi_l + \kappa^2 \Psi^2)\,\delta_{ik}. \tag{121}$$

On account of the field equation (116) this tensor is easily seen to satisfy the equation

$$\frac{\partial T_{ik}}{\partial x_k} = 0 \tag{122}$$

holding for a closed system.

The scalar field is a special case of a general field described by a number of field variables

$$Q^\zeta = Q^\zeta(x_i) = (Q^1(x_i),\, Q^2(x_i),\, ..). \tag{123}$$

Suppose that the field equations are derivable from a variational principle

$$\delta \int \mathfrak{L}\, d\Sigma = 0, \tag{124}$$

where

$$\mathfrak{L} = \mathfrak{L}(Q^\zeta, Q^\zeta_i) \tag{125}$$

is a certain algebraic invariant function of the field variables and their first derivatives

$$Q^\zeta_i = \frac{\partial Q^\zeta}{\partial x_i}. \tag{126}$$

This means that the field equations are of the form of the Euler equations

$$\frac{\partial \mathfrak{L}}{\partial Q^\zeta} - \frac{\partial}{\partial x_i}\left(\frac{\partial \mathfrak{L}}{\partial Q^\zeta_i}\right) = 0 \tag{127}$$

following from the variational principle (124). Since $\mathfrak{L}$ is supposed to be an invariant, the equations (127) will have the same form in every system of inertia.

In virtue of the field equations (127) the quantity

$$\theta_{ik} = -\sum_\zeta \frac{\partial \mathfrak{L}}{\partial Q^\zeta_k} Q^\zeta_i + \mathfrak{L}\,\delta_{ik} \tag{128}$$

is now easily seen to satisfy the divergence relation

$$\frac{\partial \theta_{ik}}{\partial x_k} = 0 \tag{129}$$

In fact, we have

$$\frac{\partial \theta_{ik}}{\partial x_k} = -\sum_\zeta \frac{\partial \mathfrak{L}}{\partial Q^\zeta_k}\frac{\partial Q^\zeta_i}{\partial x_k} - \sum_\zeta \frac{\partial}{\partial x_k}\left(\frac{\partial \mathfrak{L}}{\partial Q^\zeta_k}\right) Q^\zeta_i + \frac{\partial \mathfrak{L}}{\partial x_i}$$

$$= -\sum_\zeta \frac{\partial \mathfrak{L}}{\partial Q^\zeta_k}\frac{\partial Q^\zeta_i}{\partial x_k} - \sum_\zeta \frac{\partial \mathfrak{L}}{\partial Q^\zeta} Q^\zeta_i + \sum_\zeta \left(\frac{\partial \mathfrak{L}}{\partial Q^\zeta}\frac{\partial Q^\zeta}{\partial x_i} + \frac{\partial \mathfrak{L}}{\partial Q^\zeta_k}\frac{\partial Q^\zeta_k}{\partial x_i}\right) = 0$$

on account of the relation $\dfrac{\partial Q^\zeta_i}{\partial x_k} = \dfrac{\partial Q^\zeta_k}{\partial x_i}$ following from (126).

The quantity (128) may therefore be taken as the energy-momentum tensor of the field.

When $\mathfrak{L}$ is an invariant, θ_{ik} is easily seen to be a tensor. It is called the canonical energy-momentum tensor, its time component T_{44} being equal to $-\mathfrak{H}$, where

$$\mathfrak{H} = \sum_{\zeta} \frac{\partial \mathfrak{L}}{\partial Q_4^{\zeta}} Q_4^{\zeta} - \mathfrak{L} = \sum \frac{\partial \mathfrak{L}}{\partial \dot{Q}^{\zeta}} \dot{Q}^{\zeta} - \mathfrak{L} \tag{130}$$

is the Hamiltonian density.

In the case of the scalar field, the expression (128) for θ_{ik} reduces to the tensor T'_{ik} given by (121). In general, however, θ_{ik} will not be symmetrical and θ_{ik} will differ from the real energy-momentum tensor by a divergence-free tensor t_{ik}, so that

$$T_{ik} = \theta_{ik} + t_{ik} = T_{ki} \tag{131}$$

where
$$t_{ik} - t_{ki} = -(\theta_{ik} - \theta_{ki}), \qquad \frac{\partial t_{ik}}{\partial x_k} = 0 \tag{132}$$

Belinfante† and Rosenfeld‡ have given a general formula for the calculation of t_{ik}.

As another example we consider again the case of an electromagnetic field in vacuum treated in Chapter V. In this case the field variables Q^{ζ} are the components of the four-potential A_k, and the function $\mathfrak{L}$ is

$$\mathfrak{L} = -\tfrac{1}{4} F_{lm} F_{lm} = -\tfrac{1}{4}(A_{lm} - A_{ml})(A_{lm} - A_{ml})$$
$$= -\tfrac{1}{2}(A_{lm} A_{lm} - A_{lm} A_{ml}) \tag{133}$$

where
$$A_{lm} = \frac{\partial A_l}{\partial x_m}.$$

The Euler equations (127) then take the form of the Maxwell equations (V. 16) in vacuum

$$\frac{\partial}{\partial x_k}(A_{ki} - A_{ik}) = \frac{\partial}{\partial x_k} F_{ik} = 0 \tag{134}$$

or
$$\frac{\partial}{\partial x_i}\left(\frac{\partial A_k}{\partial x_k}\right) - \frac{\partial^2 A_i}{\partial x_k \partial x_k} = 0.$$

Together with the Lorentz condition (V. 22) which must be regarded as an accessory condition, this gives the wave equation

$$\Box A_i = 0.$$

† F. J. Belinfante, *Physica*, **6**, 887 (1939), ibid. **7**, 305 (1940).
‡ L. Rosenfeld, *Memoires de l'Acad. Roy. Belgique*, **6**, 30 (1940).

The canonical energy-momentum tensor (128) is now

$$\theta_{ik} = \sum_l F_{kl} A_{li} - \tfrac{1}{4}(F_{lm} F_{lm})\, \delta_{ik} \tag{135}$$

This tensor is not symmetrical and it deviates from the symmetrical electromagnetic energy-momentum tensor (V 106) by the term

$$t_{ik} = -A_{il} F_{kl}, \tag{136}$$

which satisfies the equation

$$\frac{\partial t_{ik}}{\partial x_k} = 0$$

on account of (134) and of the antisymmetry of the tensor F_{kl}.

VII

NON-CLOSED SYSTEMS. ELECTRODYNAMICS IN DIELECTRIC AND PARAMAGNETIC SUBSTANCES. THERMODYNAMICS

69. General properties of non-closed systems

A CLOSED system Σ may be divided in many ways into two non-closed systems $\Sigma^{(1)}$ and $\Sigma^{(2)}$ corresponding to a decomposition of the total energy-momentum tensor T_{ik} into two parts

$$T_{ik} = T_{ik}^{(1)} + T_{ik}^{(2)}. \tag{1}$$

In the case of electrically charged matter, $T_{ik}^{(1)}$ may, for instance, be the mechanical energy-momentum tensor and $T_{ik}^{(2)}$ the electromagnetic tensor. Defining a four-vector f_i by

$$f_i = -\frac{\partial T_{ik}^{(2)}}{\partial x_k},$$

we get from (1) and from (VI. 1),

$$\frac{\partial T_{ik}^{(1)}}{\partial x_k} = -\frac{\partial T_{ik}^{(2)}}{\partial x_k} = f_i \tag{2}$$

f_i is the four-force density produced by the system $\Sigma^{(2)}$ and acting on the system $\Sigma^{(1)}$. Thus the fundamental equations of a non-closed system are of the form

$$\frac{\partial T_{ik}}{\partial x_k} = f_i = -\frac{\partial S_{ik}}{\partial x_k}, \tag{2'}$$

where f_i is the force acting on the system with the tensor T_{ik}. The force acting on the system with the tensor S_{ik} is then $-f_i$.

The physical meaning of the space-time components T_{i4} and T_{4i} of the energy-momentum tensor of the non-closed system is as in (VI. 2, 3),

i.e.

$$\left.\begin{aligned} T_{4\iota} &= \frac{i}{c}S_\iota, \qquad T_{\iota 4} = icg_\iota, \qquad T_{44} = -h \\ g_i &= \frac{1}{ic}T_{i4} = \left(\mathbf{g}, \frac{i}{c}h\right) \end{aligned}\right\} \tag{3}$$

where h, $\mathbf{S}$, and $\mathbf{g}$ now denote the energy density, energy flux, and momentum density, respectively, of the non-closed system.

Instead of (VI. 4) and (VI 5) we now have

$$\frac{\partial g_\iota}{\partial t} + \frac{\partial T_{\iota\kappa}}{\partial x_\kappa} = f_\iota, \qquad \frac{\partial h}{\partial t} + \operatorname{div}\mathbf{S} = \frac{c}{i}f_4, \tag{4}$$

which represent the momentum and energy theorems for a non-closed system, by analogy with (IV 238, 236)

In view of the arbitrariness in the decomposition (1), the energy-momentum tensor of a non-closed system need, however, not necessarily be symmetric, but in any case we must have

$$T_{ik}-T_{ki} = -(S_{ik}-S_{ki}). \tag{5}$$

The total linear momentum and energy

$$G_i = \int g_i\, dV = \int \frac{T_{i4}}{ic}\, dV = \left(\mathbf{G}, \frac{i}{c}H\right)$$

of a non-closed finite system is, of course, in general not constant in time. Integrating (2′) over the whole physical space in an arbitrary system of inertia S, we get

$$\frac{d}{dt}\int \frac{T_{i4}}{ic}\, dV = \frac{d}{dt}G_i = \int f_i\, dV. \tag{6}$$

If $G_i(t)$ and $G_i'(t')$ represent the momentum and energy of a non-closed system in two different systems of inertia, the connexion between G_i and G_i' will not be given by (VI. 31′) This is already obvious from the fact that there is no unique connexion between the variables t and t' occurring as arguments in G_i and G_i'. But, even for a stationary system where G_i and G_i' are time-independent, the quantities G_i will not transform like the components of a four-vector (see § 70). This follows at once from the proof given in § 63 for the vector character of G_i in the case of a closed system For a non-closed system we would get, instead of (VI. 28),

$$\int\limits_{\Omega_1} b_k\, dV_k + \int\limits_{\Omega_2} b_k\, dV_k = \int\limits_{\Sigma} f_i\, d\Sigma, \tag{7}$$

and the right-hand side of (7) will be zero only for very special non-closed systems.

Also the angular momentum defined by

$$M_{ik} = \int (x_i g_k - x_k g_i)\, dV \tag{8}$$

will now be time-dependent. From (2′) we get, instead of (VI. 33),

$$\frac{\partial(x_i T_{kl} - x_k T_{il})}{\partial x_l} = x_i f_k - x_k f_i + T_{ki} - T_{ik}.$$

Hence, by integration over the whole physical space,

$$\frac{d}{dt}M_{ik} = \int (x_i f_k - x_k f_i + T_{ki} - T_{ik})\, dV. \tag{9}$$

Thus, in this case, the density of the moment of the forces has to be defined by

$$d_{ik} = x_i f_k - x_k f_i + T_{ki} - T_{ik} = x_i f_k - x_k f_i + S_{ik} - S_{ki}. \tag{10}$$

For a non-closed system the centre of mass loses its physical importance. Defining the coordinate vector of the centre of mass in the system of inertia S by the equation (VI. 39)

$$\mathbf{X}(S) = \frac{1}{H}\int h(\mathbf{x},t)\,\mathbf{x}\,dV = \frac{1}{G_4}\int \mathbf{x}g_4\,dV, \tag{11}$$

we get, by means of (9) with $i = \iota$, $k = 4$ and of (6) with $i = 4$, after a simple calculation,

$$\frac{d\mathbf{X}(S)}{dt} = \frac{c^2\mathbf{G}}{H} - \frac{\mathbf{X}(S)}{G_4}\int f_4\,dV + \frac{1}{G_4}\int \mathbf{x}f_4\,dV + \frac{1}{G_4}\int\left(\frac{i}{c}\mathbf{S} - ic\mathbf{g}\right)dV. \tag{12}$$

The velocity of the centre of mass is thus not equal to $c^2\mathbf{G}/H$ as for a closed system even if the energy-momentum tensor is symmetrical. This severely limits the value of the centre of mass as a representative point of the physical system.

In a closed system the *proper centre of mass* was the centre of mass in its own rest system. We may now also for a non-closed system try to define a representative point inside the system which at any time is the centre of mass in its momentary rest system, the rest systems being, of course, different at different times. A closer investigation shows, however,† that the representative point is not uniquely defined by this condition. In fact, even in a closed system there is an infinite number of points which at any time are centres of mass in their momentary rest systems. For if we imagine the disk defined in § 64 to rotate with constant angular velocity

$$\boldsymbol{\omega}^0 = -\frac{M_0 c^2}{|\mathbf{m}^0|^2}\mathbf{m}^0 \tag{13}$$

in the rest system S^0 of the proper centre of mass, any point on the rotating disk will be the centre of mass in its momentary rest system. Consider, for instance, a point p which at the time considered has the radius vector **a** reckoned from the centre of the disk. Its velocity is then

$$\mathbf{v} = (\boldsymbol{\omega}^0 \times \mathbf{a}) = -\frac{M_0 c^2}{|\mathbf{m}^0|^2}(\mathbf{m}^0 \times \mathbf{a}), \tag{14}$$

hence

$$\frac{\mathbf{m}^0 \times \mathbf{v}}{M_0 c^2} = -\frac{1}{|\mathbf{m}^0|^2}(\mathbf{m}^0 \times (\mathbf{m}^0 \times \mathbf{a})) = \mathbf{a}. \tag{15}$$

† C. Møller, *Ann. Inst. Henri Poincaré*, **11**, fasc. v, 251 (1950).

A comparison of (15) and (VI. 48) shows that the point p is centre of mass in a system S moving with the same velocity $\mathbf{v}$ relative to S^0 as the point itself, i e. any point on the rotating disk is centre of mass in its own rest system.

In the case of a closed system it was possible, however, to single out one point, the proper centre of mass, by the condition (VI. 40), which means that the total linear momentum of the physical system is zero in the rest system of the point This is not possible in the case of a non-closed system For if we apply the equation (12) in the momentary rest system of one of the representative points defined above, the left-hand side is zero and the equation (12) then shows that the linear momentum will, in general, not be zero in this system of inertia, and even if this should be the case at the moment considered, it will not be so at a later time. Thus a unique generalization of the Newtonian centre of gravity for non-closed relativistic systems is possible only for very special external forces (see § 70) There is one important exception, however, as we shall see in Chapter X, § 114 If the external forces are gravitational forces, and if the system is sufficiently small, it is always possible to define uniquely a proper centre of mass with all the properties of the Newtonian centre of gravity

70. Static non-closed systems

Let again T_{ik} be the energy-momentum tensor of the system considered but let f_i now be the four-force density *produced* by the system, then, according to (2′), we have in every system of inertia

$$f_i = -\frac{\partial T_{ik}}{\partial x_k}. \tag{16}$$

The system is called static if a system of coordinates S^0 exists in which all physical variables are time-independent and if, further,

$$\mathbf{G}^0 = \int \mathbf{g}^0\, dV^0 = \int \mathbf{S}^0\, dV^0 = 0 \tag{17}$$

The system as a whole is therefore at rest in the system S^0, and since all physical variables are time-independent in S^0, it is clear that also the centre of mass in S^0, as defined by (11), is at rest in S^0 The system considered thus represents a case in which an unambiguous generalization of the Newtonian centre of gravity is possible for a non-closed system.

As an example we may think of the electromagnetic field of charged matter at rest in a definite coordinate system S^0. The tensor T_{ik} is then the electromagnetic energy-momentum tensor S_{ik} which for a substance with $\epsilon = \mu = 1$ is given by (V. 106), f_i being then the electromagnetic four-force density acting on the charged matter.

Another simple example of a static non-closed system is a fluid contained in a vessel under the influence of the external pressure from the walls of the vessel.

To find the total energy and momentum in a system of inertia S with respect to which S^0 is moving with the constant velocity $\mathbf{u}$ we may use the transformation properties of a tensor and the expressions (VI. 87) for the transformation coefficients α_{ik}.

Integrating the equation

$$T_{ik} = T^0_{lm}\,\alpha_{li}\,\alpha_{mk} \tag{18}$$

over the whole space, we then get, using the Lorentz formula

$$dV = dV^0\sqrt{(1-u^2/c^2)}$$

and (17),

$$\left.\begin{aligned} \mathbf{G} &= \int \mathbf{g}\,dV = \frac{\mathbf{u}[H^0+(1/u^2)(\mathbf{u}\,.\int \mathbf{T}^0\,dV^0\,.\,\mathbf{u})\{(1-\sqrt{(1-u^2/c^2)}\}]}{c^2\sqrt{(1-u^2/c^2)}} \\ &\qquad\qquad + \frac{1}{c^2}\left(\int \mathbf{T}^0\,dV^0\,.\,\mathbf{u}\right) \\ H &= \int h\,dV = \frac{H^0+(1/c^2)(\mathbf{u}\,.\int \mathbf{T}^0\,dV^0\,.\,\mathbf{u})}{\sqrt{(1-u^2/c^2)}} \end{aligned}\right\}, \tag{19}$$

where $\mathbf{T}^0$ is the spatial tensor with the components $T^0_{\iota\kappa}$. Although $\mathbf{G}$ and H are constant in time they do not transform like the components of a four-vector. In general, this may be taken as a proof that the system considered is non-closed. For an elastic body, the equations (19) are obtained from (VI. 92, 89) by integration, if we assume that the velocity $\mathbf{u}$ is constant throughout the body. Such a system cannot therefore be closed unless the stress tensor $\mathbf{t}^0 = \mathbf{T}^0 = 0$ everywhere in the body.

If $\mathbf{T}^0$ is of the form

$$T^0_{\iota\kappa} = p^0\,\delta_{\iota\kappa}, \tag{20}$$

(19) reduces to the same equations as for a perfect fluid, i.e. (VI. 112).

71. Electrostatic systems. Classical models of the electron

Let us now consider in a little more detail the case of charged matter at rest in a system of coordinates S^0. If $\epsilon = \mu = 1$, the tensor S_{ik} is given by (V. 106, 114). Since the field is electrostatic in S^0 we have $\mathbf{H}^0 = 0$, and $\mathbf{E}^0$ is constant in time, i.e.

$$h^0 = \tfrac{1}{2}|\mathbf{E}^0|^2, \qquad \mathbf{S}^0 = \mathbf{g}^0 = 0, \qquad S^0_{\iota\kappa} = -E^0_\iota E^0_\kappa + \tfrac{1}{2}|E^0|^2\,\delta_{\iota\kappa}. \tag{21}$$

Let us in particular consider the case of a spherically symmetrical distribution of the electric charges. In this case the field will also be spherically symmetrical, $\mathbf{E}^0$ being directed along the radius vector connecting the centre of the charge distribution with the point considered. Hence

$$\int E_\iota^0 E_\kappa^0\, dV^0 = \tfrac{1}{3}\int |\mathbf{E}^0|^2\, dV^0\, \delta_{\iota\kappa}$$

and $$\int S_{\iota\kappa}^0\, dV^0 = \tfrac{1}{6}\int |\mathbf{E}^0|^2\, dV^0\, \delta_{\iota\kappa} = \tfrac{1}{3}\int h^0\, dV^0\, \delta_{\iota\kappa} = \tfrac{1}{3}H^0\, \delta_{\iota\kappa} \qquad (22)$$

Using (22) in (19) we get for the total electromagnetic momentum and energy of a spherically symmetric charge distribution

$$\mathbf{G}_{\mathrm{el}} = \frac{4}{3}\frac{H_{\mathrm{el}}^0}{c^2}\frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}}, \qquad H_{\mathrm{el}} = H_{\mathrm{el}}^0\frac{1+\frac{1}{3}u^2/c^2}{\sqrt{(1-u^2/c^2)}}. \qquad (23)$$

Such a system represents a classical model of the electron, the fundamental equations of Lorentz's electron theory being identical with Maxwell's equations for substances with $\epsilon = \mu = 1$. Lorentz put forward the idea that the mass, energy, and momentum of the electron could be of purely electromagnetic origin, but from (23) we see that this is impossible,† since the dependence of the electromagnetic energy on the velocity differs from the relativistic formula (III 31) for the energy of a particle. Since the quantities $(\mathbf{G}_{\mathrm{el}}, (i/c)H_{\mathrm{el}})$ do not transform like the components of a four-vector, we have to deal with a typically non-closed system, and in order to get a consistent classical picture of the electron we must assume the existence of non-electromagnetic energies and momenta inside the electron at least as long as Maxwell's equations are supposed to hold throughout the whole space.

Let us now assume that the charge e is uniformly distributed over the surface of an elastic sphere of radius a in the rest system. If $\mathbf{n}$ is a unit vector in the direction of the radius vector, the solution of Maxwell's equations is

$$\mathbf{E}^0 = \frac{e}{4\pi r^2}\mathbf{n} \text{ for } r > a, \qquad \mathbf{E}^0 = 0 \text{ for } r < a, \qquad \mathbf{H}^0 = 0, \qquad (24)$$

where r is the distance from the centre of the sphere. Thus we get from (21)

$$\left.\begin{aligned} S_{\iota\kappa}^0 &= -\frac{e^2}{(4\pi)^2 r^4}n_\iota n_\kappa + \frac{1}{2}\frac{e^2}{(4\pi)^2 r^4}\delta_{\iota\kappa}, \quad \text{for } r > 0 \\ H_{\mathrm{el}}^0 &= \tfrac{1}{2}\int |\mathbf{E}^0|^2\, dV^0 = \frac{e^2}{(4\pi)^2 2}4\pi\int_a^\infty \frac{r^2\, dr}{r^4} = \frac{e^2}{8\pi a} = m_0^{\mathrm{el}}c^2 \end{aligned}\right\}. \qquad (25)$$

† M Abraham, *Phys. ZS* 5, 576 (1904)

Here the charge e is measured in Heaviside units and m_0^{el} is the electromagnetic contribution to the rest mass of the particle.

According to (V 109) the electric force per unit surface on the sphere is

$$t_\iota^{\mathrm{el}} = -S_{\iota\kappa}^0 n_\kappa = \frac{e^2}{2(4\pi)^2 a^4} n_\iota, \tag{26}$$

which must be in equilibrium with the elastic force Thus the elastic stress tensor inside the sphere must be of the form

$$t_{\iota\kappa}^0 = p^0 \, \delta_{\iota\kappa}, \tag{27}$$

where
$$p^0 = -\frac{e^2}{2(4\pi)^2 a^4} = -\frac{H_{\mathrm{el}}^0}{4\pi a^3} \tag{28}$$

The total mechanical energy and momentum can now be obtained from (VI 112)

$$\left.\begin{aligned} \mathbf{G}_{\mathrm{me}} &= \frac{H_{\mathrm{me}}^0 + p^0 V^0}{c^2} \frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}} = \left(m_0^{\mathrm{me}} - \frac{1}{3}\frac{H_{\mathrm{el}}^0}{c^2}\right)\frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}} \\ H_{\mathrm{me}} &= \frac{H_{\mathrm{me}}^0 + (p^0 V^0/c^2)u^2}{\sqrt{(1-u^2/c^2)}} = \frac{m_0^{\mathrm{me}} c^2 - \frac{1}{3}(H_{\mathrm{el}}^0/c^2)u^2}{\sqrt{(1-u^2/c^2)}} \end{aligned}\right\} \tag{29}$$

Adding the expressions (23) and (29) we get for the total energy and momentum

$$\left.\begin{aligned} \mathbf{G} &= \mathbf{G}_{\mathrm{me}} + \mathbf{G}_{\mathrm{el}} = \frac{(H_{\mathrm{me}}^0 + H_{\mathrm{el}}^0)}{c^2} \frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}} = \frac{H^0}{c^2} \frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}} \\ H &= H_{\mathrm{me}} + H_{\mathrm{el}} = \frac{(H_{\mathrm{me}}^0 + H_{\mathrm{el}}^0)}{\sqrt{(1-u^2/c^2)}} = \frac{H^0}{\sqrt{(1-u^2/c^2)}} \end{aligned}\right\}, \tag{30}$$

as we should have for a closed system. A system of that kind was used for the first time by Poincaré as a model of the electron † Poincaré did not specify the nature of the forces which in his model counterbalance the electric forces in the electron; he simply assumed the existence of such forces of non-electromagnetic nature and a corresponding energy-momentum tensor which together with the electromagnetic tensor defines a total energy-momentum tensor T_{ik} satisfying the condition $\partial T_{ik}/\partial x_k = 0$, characteristic of a closed system

In contradistinction to this dualistic point of view, which requires the introduction of field quantities of a non-electromagnetic nature, Mie‡ and Born§ advocated a unitary point of view in which only electromagnetic field variables are introduced These field variables must then satisfy equations which deviate from the Maxwell equations inside the

† H. Poincaré, *Rend Pal.* **21**, 129 (1906)
‡ G Mie, *Ann d. Phys* **37**, 511 (1912), **39**, 1 (1912), **40**, 1 (1913)
§ M Born, *Proc Roy Soc* A, **143**, 410 (1934)

electron where the field is strong These field equations are non-linear and the corresponding energy-momentum tensor S_{ik} satisfies the necessary condition $\partial S_{ik}/\partial x_k = 0$, i e. the self-force $f_i = -\partial S_{ik}/\partial x_k$ is zero.

The final solution of the problem of the electron and of the other elementary particles can probably not be found on a classical basis. Besides the introduction of Planck's quantum of action it may even be necessary to introduce a new fundamental constant of the dimension of a length † But the above considerations show that, as long as one assumes the existence of an energy-momentum tensor of the system, the theory of relativity requires the vanishing of the self-force, i e of the four-dimensional divergence of this tensor

72. The fundamental equations of electrodynamics in stationary matter

As shown by Lorentz,‡ Maxwell's phenomenological equations of electrodynamics for stationary matter may be derived from the fundamental equations of the electron theory by averaging over regions in space which are small from the macroscopic point of view, but still so large that they contain a large number of electrons. Since the equations (V 13, 16) of the electron theory are covariant in form, it must be possible also to find the 'macroscopic' equations of electrodynamics in *moving* bodies by averaging over appropriate small space-time regions This was actually done by Born§ and Dallenbach ||

However, if we assume the validity of Maxwell's phenomenological equations for a body at rest, it is possible to find the corresponding equations in moving bodies simply by performing a Lorentz transformation This method was used for the first time by Minkowski.†† The principle of relativity requires that Maxwell's equations for stationary matter must hold in that system of coordinates S^0 in which the matter is at rest, irrespective of the velocity of this system with respect to the fixed stars. Thus we have in S^0

$$\left.\begin{aligned} \operatorname{curl}^0 \mathbf{E}^0 + \frac{1}{c}\frac{\partial \mathbf{B}^0}{\partial t^0} &= 0, \qquad & \operatorname{div}^0 \mathbf{B}^0 &= 0 \\ \operatorname{curl}^0 \mathbf{H}^0 - \frac{1}{c}\dot{\mathbf{D}}^0 &= \mathbf{J}^0/c, \qquad & \operatorname{div} \mathbf{D}^0 &= \rho^0 \end{aligned}\right\}, \tag{31}$$

† W Heisenberg, *Ann d Phys* **32**, 20 (1938)

‡ See ref, Chap I, p 21.

§ Minkowski-Born, *Math Ann* **68**, 526 (1910), see also A D Fokker, *Phil. Mag.* **39**, 404 (1920)

|| W Dallenbach, Diss, Zurich, 1918, *Ann d Phys* **58**, 523 (1919)

†† Minkowski, H, *Gott Nachr*, p 53 (1908), *Math Ann* **68**, 472 (1910)

where $\mathbf{E}^0$, $\mathbf{D}^0$, $\mathbf{H}^0$, $\mathbf{B}^0$ denote the electric field strength, electric displacement, magnetic field strength, and magnetic induction, respectively. ρ^0 and $\mathbf{J}^0$ are the macroscopic charge and current densities. All these quantities may in principle be determined by means of macroscopic experiments in S^0. $\mathbf{E}^0$ and $\mathbf{D}^0$, for instance, are defined as the forces on a small test body of unit charge inserted at the point considered into small crevasses cut in the matter parallel or perpendicular to the field, respectively. Similarly, $\mathbf{H}^0$ and $\mathbf{B}^0$ are the corresponding forces on a magnetic test body of unit magnetic pole strength.

Besides the field equations (31) we have in isotropic dielectric and paramagnetic substances the following constitutive equations connecting the field variables with the constitution of the matter

$$\mathbf{D}^0 = \epsilon\mathbf{E}^0, \qquad \mathbf{B}^0 = \mu\mathbf{H}^0, \qquad \mathbf{J}^0 = \sigma\mathbf{E}^0, \tag{32}$$

where ϵ is the dielectric constant, μ the magnetic permeability, and σ the electrical conductivity. The last equation is the mathematical expression of Ohm's law.

73. Minkowski's field equations in uniformly moving bodies

Consider two antisymmetrical tensors F_{ik} and H_{ik}. According to (IV. 80, 80′) the tensor F_{ik} defines a pair of space vectors $\mathbf{B}$ and $\mathbf{E}$ in the arbitrary system of coordinates S by the equations

$$\mathbf{B} = (F_{23}, F_{31}, F_{12}), \qquad i\mathbf{E} = (F_{41}, F_{42}, F_{43}) \tag{33}$$

$\mathbf{B}$ is an axial vector and $\mathbf{E}$ a polar vector. Similarly, the tensor H_{ik} defines a polar vector $\mathbf{D}$ and an axial vector $\mathbf{H}$ by the equations

$$\mathbf{H} = (H_{23}, H_{31}, H_{12}), \qquad i\mathbf{D} = (H_{41}, H_{42}, H_{43}) \tag{34}$$

Further, consider a four-vector with the components

$$J_i = (\mathbf{J}/c, i\rho) \tag{35}$$

in the system S. When the components of the tensors F_{ik}, H_{ik}, and J_i are given in one system of coordinates, we can calculate the components of these tensors in any other system by means of the transformation equations (IV. 81′) and (IV. 29) of antisymmetrical tensors and vectors. For the components of J_i we thus get

$$\left.\begin{aligned} \mathbf{J} &= \mathbf{J}' + \frac{\mathbf{v}}{v^2}\,\frac{(\mathbf{v}\cdot\mathbf{J}')\left(1-\sqrt{(1-v^2/c^2)}\right)+\rho' v^2}{\sqrt{(1-v^2/c^2)}} \\ \rho &= \frac{\rho' + (1/c^2)(\mathbf{v}\cdot\mathbf{J}')}{\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}, \tag{36}$$

where v is the velocity of S' relative to S.

If we now define the tensors F_{ik}, H_{ik}, J_i so that the quantities **E**, **D**, **H**, **B**, **J**, ρ are identical with the macroscopic electromagnetic variables $\mathbf{E}^0$, $\mathbf{D}^0$, $\mathbf{H}^0$, $\mathbf{B}^0$, $\mathbf{J}^0$, ρ^0 in the rest system S^0 of the matter, the field equations of electrodynamics in any system of coordinates must have the form

$$\frac{\partial F_{ik}}{\partial x_l}+\frac{\partial F_{kl}}{\partial x_i}+\frac{\partial F_{li}}{\partial x_k}=0, \tag{37 a}$$

$$\frac{\partial H_{ik}}{\partial x_k}=J_i \tag{37 b}$$

For in the rest system S^0 the equations (37) are then identical with Maxwell's equations (31) and, since they are tensor equations, they must hold in any system of inertia On account of (IV 187) the equations (37 a) may also be written

$$\frac{\partial F^*_{ik}}{\partial x_k}=0, \tag{37 a'}$$

where F^*_{ik} is the pseudo-tensor dual to F_{ik}

Using (33), (34), (35) in (37) we thus get in every system of coordinates

$$\operatorname{curl}\mathbf{E}+\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}=0, \qquad \operatorname{div}\mathbf{B}=0, \tag{38 a}$$

$$\operatorname{curl}\mathbf{H}-\frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}=\frac{\mathbf{J}}{c}, \qquad \operatorname{div}\mathbf{D}=\rho \tag{38 b}$$

The quantities $J_i=(\mathbf{J}/c, i\rho)$ may be interpreted as current and charge densities in the system S, for from (37 b) we get the continuity equation

$$\frac{\partial J_i}{\partial x_i}=\frac{\partial^2 H_{ik}}{\partial x_i\,\partial x_k}=0 \tag{39}$$

Further, in an insulator, where $\mathbf{J}^0=0$ in the rest system, we get from the equations (36)

$$\rho=\frac{\rho^0}{\sqrt{(1-u^2/c^2)}}, \qquad \mathbf{J}=\frac{\rho^0\mathbf{u}}{\sqrt{(1-u^2/c^2)}}=\rho\mathbf{u}, \tag{40}$$

where **u** is the velocity of the ponderable matter relative to S Hence

$$\delta e=\rho\,\delta V=\rho^0\,\delta V^0, \tag{41}$$

i e the charge contained in an infinitesimal piece of matter of volume $\delta V=\delta V^0\sqrt{(1-u^2/c^2)}$ is invariant and the current **J** is a pure convection current

In the general case, the current density can always be written as a sum of the convection current $\rho\mathbf{u}$ and the conduction current **C**

$$\mathbf{J}=\rho\mathbf{u}+\mathbf{C}, \tag{42}$$

but this separation is not relativistically invariant. Even if $\rho^0 = 0$ in the rest system, so that $\mathbf{J}^0 = \mathbf{C}^0$ is a pure conduction current in this system, the charge density ρ in S will be different from zero, which means that we also have a convection current $\rho\mathbf{u} \neq 0$ in S. In fact, we get from (36) in this case

$$\left.\begin{aligned} \mathbf{J} &= \mathbf{C}^0 + \frac{\mathbf{u}}{u^2}\frac{(\mathbf{u}.\mathbf{C}^0)\{1-\sqrt{(1-u^2/c^2)}\}}{\sqrt{(1-u^2/c^2)}} = \rho\mathbf{u}+\mathbf{C} \\ \rho &= \frac{(\mathbf{u}.\mathbf{C}^0)}{c^2\sqrt{(1-u^2/c^2)}}, \qquad \mathbf{C} = \mathbf{C}^0 + \frac{\mathbf{u}(\mathbf{u}.\mathbf{C}^0)}{u^2}\{\sqrt{(1-u^2/c^2)}-1\} \end{aligned}\right\} \tag{43}$$

We can, however, make a relativistically invariant decomposition of J_i by writing

$$J_i = \frac{\rho^0 U_i}{c} + s_i, \tag{44}$$

where ρ^0 is the (invariant) charge density in the rest system S^0 of the matter, and

$$U_i = \left\{\frac{\mathbf{u}}{\sqrt{(1-u^2/c^2)}}, \frac{ic}{\sqrt{(1-u^2/c^2)}}\right\}$$

is the four-velocity of the ponderable matter. $s_i = (\mathbf{s}, s_4)$ is a four-vector with the components

$$s_i^0 = (\mathbf{J}^0/c, 0) \tag{45}$$

in S^0. Thus

$$\left.\begin{aligned} s_i U_i = s_i^0 U_i^0 = 0, \qquad J_i U_i = -c\rho^0 \\ s_i = J_i + (1/c^2)(J_l U_l)U_i \end{aligned}\right\} \tag{46}$$

While J_i has a direct physical meaning, the field variables **E**, **D**, **H**, **B**, occurring in the field equations (38) have no simple physical significance in contrast to the field variables $\mathbf{E}^0$, $\mathbf{D}^0$, $\mathbf{H}^0$, $\mathbf{B}^0$ in the rest system, which could be determined by simple macroscopic experiments. So far they are defined only by the transformation equations by which they may be expressed in terms of the field variables in the rest system S^0.

Let us now consider the four-vector F_i defined by

$$F_i = \frac{1}{c}F_{ik}U_k \tag{47}$$

From (33) we get for the components of F_i in S

$$F_i = \left(\frac{\mathbf{E}+(1/c)(\mathbf{u}\times\mathbf{B})}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{E}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right). \tag{48}$$

Its components in the rest system are thus

$$F_i^0 = (\mathbf{E}^0, 0), \tag{49}$$

i e F_i is the four-force acting on a unit charge placed at rest relative to the matter in a longitudinal crevasse cut in the matter If we put

$$\tilde{\mathbf{E}} = \mathbf{E}+\frac{1}{c}(\mathbf{u}\times\mathbf{B}), \tag{50}$$

(48) may be written

$$F_i = \left(\frac{\tilde{\mathbf{E}}}{\sqrt{(1-u^2/c^2)}}, \frac{i(\tilde{\mathbf{E}}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right), \tag{51}$$

and a comparison with (IV. 54) shows that $\tilde{\mathbf{E}}$ is the force on the test body in question measured in the system S.

Similarly, the four-vector

$$K_i = \frac{1}{c}H_{ik}U_k = \left(\frac{\mathbf{D}+(1/c)(\mathbf{u}\times\mathbf{H})}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{D}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right)$$
$$= \left(\frac{\tilde{\mathbf{D}}}{\sqrt{(1-u^2/c^2)}}, \frac{i(\tilde{\mathbf{D}}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right), \tag{52}$$

with the components $K_i^0 = (\mathbf{D}^0, 0)$ in the rest system, represents the four-force on a unit test body inserted at rest in a transverse crevasse Thus also

$$\tilde{\mathbf{D}} = \mathbf{D}+\frac{1}{c}(\mathbf{u}\times\mathbf{H}) \tag{53}$$

has a simple physical meaning. it represents the force on the test body in question measured in the system S.

Further, if F^*_{ik} and H^*_{ik} are the pseudo-tensors dual to F_{ik} and H_{ik} (see IV 108, 109), we can form the two pseudo-vectors

$$F^*_i = \frac{1}{c}F^*_{ik}U_k = \left(\frac{\mathbf{B}-(\mathbf{u}\times\mathbf{E})/c}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{B}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right)$$
$$= \left(\frac{\tilde{\mathbf{B}}}{\sqrt{(1-u^2/c^2)}}, \frac{i(\tilde{\mathbf{B}}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right), \tag{54}$$

$$K^*_i = \frac{1}{c}H^*_{ik}U_k = \left(\frac{\mathbf{H}-(\mathbf{u}\times\mathbf{D})/c}{\sqrt{(1-u^2/c^2)}}, \frac{i(\mathbf{H}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right)$$
$$= \left(\frac{\tilde{\mathbf{H}}}{\sqrt{(1-u^2/c^2)}}, \frac{i(\tilde{\mathbf{H}}.\mathbf{u})/c}{\sqrt{(1-u^2/c^2)}}\right). \tag{55}$$

F^*_i and K^*_i are obviously the four-forces acting on a unit magnetic pole placed at rest relative to the matter in transversal and longitudinal crevasses, respectively, as is seen at once when one considers the components of these pseudo-vectors in the rest system. Thus

$$\tilde{\mathbf{B}} = \mathbf{B} - \frac{\mathbf{u}\times\mathbf{E}}{c}, \tag{56}$$

$$\tilde{\mathbf{H}} = \mathbf{H} - \frac{\mathbf{u}\times\mathbf{D}}{c} \tag{57}$$

are the forces on these unit magnetic poles measured in the system S.

The vectors $\tilde{\mathbf{E}}$, $\tilde{\mathbf{D}}$, $\tilde{\mathbf{H}}$, $\tilde{\mathbf{B}}$ (or the four-vectors F_i, K_i, F_i^*, K_i^*) may thus in principle be obtained by direct physical measurements performed by an observer in S. By means of (50), (53), (56), and (57) we can now also express F_{ik} and H_{ik} in terms of the quantities F_i, K_i, F_i^*, K_i^*. We get

$$\left.\begin{aligned} F_{ik} &= \frac{1}{c}(U_i F_k - U_k F_i) + \frac{1}{ic}\delta_{iklm} F_l^* U_m \\ H_{ik} &= \frac{1}{c}(U_i K_k - U_k K_i) + \frac{1}{ic}\delta_{iklm} K_l^* U_m \end{aligned}\right\}, \tag{58}$$

where δ_{iklm} is the Levi-Civita symbol defined in § 43 The equations (58) are easily seen to be true in the rest system S^0 and, since both sides of the equations transform like tensors, they must hold generally

Since the vector $\mathbf{u}$ is a constant we have

$$\operatorname{curl}(\mathbf{u}\times\mathbf{B}) = -(\mathbf{u}\operatorname{grad})\mathbf{B} + \mathbf{u}\operatorname{div}\mathbf{B} = -(\mathbf{u}\operatorname{grad})\mathbf{B},$$

$$\operatorname{curl}(\mathbf{u}\times\mathbf{D}) = -(\mathbf{u}\operatorname{grad})\mathbf{D} + \mathbf{u}\operatorname{div}\mathbf{D} = -(\mathbf{u}\operatorname{grad})\mathbf{D} + \rho\mathbf{u}$$

Therefore the field equations (38) can also be written in the form

$$\operatorname{curl}\tilde{\mathbf{E}} + \frac{1}{c}\frac{d\mathbf{B}}{dt} = 0, \qquad \operatorname{curl}\tilde{\mathbf{H}} - \frac{1}{c}\frac{d\mathbf{D}}{dt} = \mathbf{C}/c, \tag{59}$$

$$\operatorname{div}\mathbf{B} = 0, \qquad \operatorname{div}\mathbf{D} = \rho, \tag{60}$$

where $\qquad \dfrac{d\mathbf{B}}{dt} = \dfrac{\partial\mathbf{B}}{\partial t} + (\mathbf{u}\operatorname{grad})\mathbf{B}, \qquad \dfrac{d\mathbf{D}}{dt} = \dfrac{\partial\mathbf{D}}{\partial t} + (\mathbf{u}\operatorname{grad})\mathbf{D}$

are the substantial time derivatives of $\mathbf{B}$ and $\mathbf{D}$, and $\mathbf{C}$ is the conduction current defined by (42)

So far we have considered only one material substance moving with constant velocity $\mathbf{u}$ However, since the field equations are linear, the fields are additive and the equations (37) must hold also in the case of several bodies, separated by a vacuum, moving uniformly with different velocities The field equations (37) will, however, represent a good approximation for a system of moving bodies only as long as the *accelerations* of the bodies due to the electromagnetic forces can be regarded as small.

74. The constitutive equations in four-dimensional language. Boundary conditions

According to the first two equations of the set (32) the forces on a unit test body placed in transversal and longitudinal crevasses are proportional, the constants of proportionality being ϵ and μ according as the test body is an electric or a magnetic pole. Therefore we must have

$$\tilde{\mathbf{D}} = \epsilon\tilde{\mathbf{E}}, \qquad \tilde{\mathbf{B}} = \mu\tilde{\mathbf{H}} \tag{61}$$

or

$$K_i = \epsilon F_i, \qquad F_i^* = \mu K_i^*. \tag{62}$$

These equations may also be written

$$H_{ik} U_k = \epsilon F_{ik} U_k, \tag{63a}$$

$$F^*_{ik} U_k = \mu H^*_{ik} U_k, \tag{63b}$$

the last equation being identical with the tensor equation

$$F_{ik} U_l + F_{kl} U_i + F_{li} U_k = \mu(H_{ik} U_l + H_{kl} U_i + H_{li} U_k) \tag{63c}$$

In the rest system equations (61), (62) or (63) reduce to the first two equations of (32). The last equation of the set (32), Ohm's law, may be put into the form

$$s_i = \sigma F_i/c, \tag{64}$$

as is seen from (45) and (49) when the vector equation (64) is written down in the rest system. Since $s_i U_i = F_i U_i = 0$, only the first three equations (64) are independent, i.e. (64) is equivalent to

$$\mathbf{s} = \frac{\sigma\tilde{\mathbf{E}}/c}{\sqrt{(1-u^2/c^2)}} \tag{65}$$

On account of (46) and (47) Ohm's law can also be written

$$J_i + \frac{(J_k U_k)}{c^2} U_i = \frac{\sigma}{c^2} F_{ik} U_k. \tag{66}$$

The field equations (37) together with the constitutive equations (63) and (66) enable us to determine the field when the charge and current distributions are known.

At the boundary between the ponderable matter and the vacuum the tangential components of $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{H}}$ must be continuous, as is seen in the usual way from (59) by integrating these equations over infinitesimal surfaces bounded by a small rectangle with two opposite sides immediately inside and outside the boundary of the matter. It is here understood that the $\mathbf{u}$ occurring in the definition of $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{H}}$ is put equal to the velocity of the matter also outside the boundary. Further, we find by integration of (60) over a small cylinder with end surfaces immediately outside and inside the boundary that the normal component of **B** must be continuous at the boundary, while the change ΔD_n in the normal component of **D** is equal to the surface density of charge on the boundary

75. Electromagnetic energy-momentum tensor and four-force density

In Chapter V we have seen that the four-force density in the electron theory is given by

$$f_i = F_{ik} s_k,$$

where s_i is the current density of this theory. This expression followed immediately from the observation that by the very definition of the electric field strength the force density in the rest system of the charge is $\rho^0 \mathbf{E}^0$. In ponderable matter with ϵ and μ different from 1 it is not so easy to find a unique expression for the force density acting on the matter In the first place, we have in general a conduction current in the rest system of the matter, and even in an insulator it is not evident that the force in the rest system is $\rho^0 \mathbf{E}^0$ because the electric field strength is defined as the force on a unit test body placed in a crevasse cut in the matter This uncertainty in the definition of the force density gives rise to a corresponding uncertainty in the definition of the electromagnetic energy-momentum tensor.

However, let us consider the four-vector $F_{il} J_l$ which is the analogue of the four-force density in the electron theory From the field equations (37) we get

$$F_{il} J_l = F_{il} \frac{\partial H_{lk}}{\partial x_k} = \frac{\partial (F_{il} H_{lk})}{\partial x_k} - \frac{\partial F_{il}}{\partial x_k} H_{lk}$$

$$= -\frac{\partial (F_{il} H_{kl})}{\partial x_k} - \frac{1}{2}\left(\frac{\partial F_{il}}{\partial x_k} + \frac{\partial F_{ki}}{\partial x_l}\right) H_{lk}$$

$$= -\frac{\partial (F_{il} H_{kl})}{\partial x_k} + \frac{1}{2} \frac{\partial F_{lk}}{\partial x_i} H_{lk}$$

$$= -\frac{\partial}{\partial x_k} (F_{il} H_{kl}) + \frac{1}{4} \frac{\partial}{\partial x_i} (F_{lk} H_{lk}) + \frac{1}{4} \left[\frac{\partial F_{lk}}{\partial x_i} H_{lk} - F_{lk} \frac{\partial H_{lk}}{\partial x_i}\right].$$

Hence

$$F_{il} J_l + \frac{1}{4}\left[F_{kl} \frac{\partial H_{kl}}{\partial x_i} - \frac{\partial F_{kl}}{\partial x_i} H_{kl}\right] = -\frac{\partial S_{ik}}{\partial x_k}, \tag{67}$$

where

$$S_{ik} = F_{il} H_{kl} - \tfrac{1}{4}(F_{lm} H_{lm})\, \delta_{ik}. \tag{68}$$

From (33) and (34) we get for the components of this tensor

$$\left.\begin{aligned} S_{\iota\kappa} &= -t_{\iota\kappa}, \\ t_{\iota\kappa} &= E_\iota D_\kappa + H_\iota B_\kappa - \tfrac{1}{2}(\mathbf{E}\cdot\mathbf{D} + \mathbf{H}\cdot\mathbf{B})\, \delta_{\iota\kappa} \end{aligned}\right\} \tag{69}$$

where $t_{\iota\kappa}$ in the rest system is identical with Maxwell's stress tensor in ponderable matter.

Further

$$\left.\begin{aligned}(S_{41}, S_{42}, S_{43}) &= \frac{i}{c}\mathbf{S}\\ \text{where}\qquad \mathbf{S} &= c(\mathbf{E}\times\mathbf{H})\end{aligned}\right\}, \tag{70}$$

is Poynting's vector, and

$$S_{44} = -h, \qquad h = \tfrac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B}). \tag{71}$$

In the rest system $\mathbf{S}$ and h are identical with the usually recognized expressions for the electromagnetic energy flux and energy density in stationary matter

Finally,

$$\left.\begin{aligned}(S_{14}, S_{24}, S_{34}) &= ic\mathbf{g}\\ \text{where}\qquad \mathbf{g} &= \frac{1}{c}(\mathbf{D}\times\mathbf{B})\end{aligned}\right\} \tag{72}$$

The form of the equations (67) suggests that the left-hand side of (67) is the electromagnetic four-force density f_i and that S_{ik} as given by (68) represents the electromagnetic energy-momentum tensor This would mean that the quantities $\mathbf{S}$, h, $\mathbf{g}$ in every system of coordinates should be interpreted as the electromagnetic energy flux, energy density, and momentum density, respectively The above expressions for $\mathbf{S}$, h, and $\mathbf{g}$ are due to Minkowski,† in the case when $\epsilon = \mu = 1$ they reduce to the corresponding expressions of the electron theory.

If we confine ourselves to homogeneous and isotropic bodies it is easily seen that the second term on the left-hand side of (67) is zero In the rest system this term is

$$\frac{1}{4}\left[F^0_{kl}\frac{\partial H^0_{kl}}{\partial x^0_i}-\frac{\partial F^0_{kl}}{\partial x^0_i}H^0_{kl}\right] = \frac{1}{2}\left[\mathbf{B}^0.\frac{\partial\mathbf{H}^0}{\partial x^0_i}-\mathbf{E}^0\frac{\partial\mathbf{D}^0}{\partial x_i}-\frac{\partial\mathbf{B}^0}{\partial x^0_i}.\mathbf{H}^0+\frac{\partial\mathbf{E}^0}{\partial x^0_i}.\mathbf{D}^0\right]$$

$$= -\frac{1}{2}\left[|\mathbf{H}^0|^2\frac{\partial\mu}{\partial x^0_i}+|\mathbf{E}^0|^2\frac{\partial\epsilon}{\partial x^0_i}\right]$$

on account of (32). Thus, if ϵ and μ are constant, this term is zero in S^0, but a vector with zero components in one system of coordinates is zero in all systems Hence, inside a homogeneous and isotropic body we have

$$f_i = F_{il}J_l = \left(\rho\mathbf{E}+\frac{1}{c}(\mathbf{J}\times\mathbf{B}), \frac{i}{c}(\mathbf{E}.\mathbf{J})\right), \tag{73}$$

i e

$$\left.\begin{aligned}\mathbf{f} &= \rho\mathbf{E}+\frac{1}{c}(\mathbf{J}\times\mathbf{B}) = \rho\left(\mathbf{E}+\frac{1}{c}(\mathbf{u}\times\mathbf{B})\right)+\frac{1}{c}(\mathbf{C}\times\mathbf{B})\\ f_4 &= \frac{i}{c}(\mathbf{E}.\mathbf{J}) = \frac{i}{c}\mathbf{E}(\rho\mathbf{u}+\mathbf{C}) = \frac{i}{c}[\mathbf{f}.\mathbf{u}+\tilde{\mathbf{E}}.\mathbf{C}]\end{aligned}\right\}. \tag{74}$$

† See ref p 195

In the rest system this expression for f_4 is in accordance with Joule's expression for the heat q^0 developed in the body per unit time and volume. In fact, we have

$$f_4^0 = \frac{i}{c}(\mathbf{E}^0 . \mathbf{J}^0) = \frac{i}{c}(\mathbf{E}^0\ \mathbf{C}^0) = \frac{i}{c}q^0 \tag{75}$$

In an arbitrary system of coordinates $(\mathbf{f}.\mathbf{u})$ is the mechanical work, and $q = (\tilde{\mathbf{E}}\ \mathbf{C})$ thus must represent the heat production in accordance with (IV 217) The four equations

$$f_i = -\frac{\partial S_{ik}}{\partial x_k} \tag{76}$$

represent in the usual way the momentum and energy laws

From (73) we get

$$f_i U_i = -U_i F_{il} J_l = U_i^0 F_{il}^0 J_l^0 = -(\mathbf{E}^0 . \mathbf{J}^0) = -q^0. \tag{77}$$

According to the considerations in § 50, we are here dealing with a typical example of forces which produce a change in the total proper mass of the matter

Minkowski's electromagnetic energy-momentum tensor satisfies the same identity

$$S_{ii} = F_{il} H_{il} \quad (F_{kl} H_{kl}) = 0 \tag{78}$$

as the tensor of the electron theory, but it is not symmetrical, i e

$$S_{ik} \neq S_{ki} \tag{79}$$

In the rest system S^0 the space part (69) is symmetric for an isotropic body on account of (32), but for the mixed space-time component we have in S^0

$$S_{i4}^0 - S_{4i}^0 = ic\left(g_i^0 - \frac{S_i^0}{c^2}\right) = i(\epsilon\mu - 1)(\mathbf{E}^0 \times \mathbf{H}^0)_i \neq 0 \tag{80}$$

In any other system of reference we therefore also have $S_{\iota\kappa} \neq S_{\kappa\iota}$ even in an isotropic body

This non-symmetry of Minkowski's energy-momentum tensor has given rise to a long discussion in the literature † It was generally felt that this property represented a real difficulty for Minkowski's theory Abraham‡ therefore tried to construct a symmetrical expression for the electromagnetic energy-momentum tensor In the rest system S^0 Abraham's tensor agrees with (69), (70), and (71) at least for isotropic

† M Abraham, *Rend Pal* **28** (1909), *Ann d Phys* **44**, 537 (1914), W Dallenbach, op cit (p 195), M von Laue, *Die Relativitatstheorie*, vol I, 3rd ed , § 24, Braunschweig, 1919, W Pauli, *Encykl d Math Wiss* vol 2 (1920), p 667, R C Tolman, *Relativity, Thermodynamics, and Cosmology*, § 54, Oxford, 1934, Jg Tamm, *Journ of Phys USSR*, **1**, 439 (1939)

‡ M Abraham, *Rend Pal* **28** (1909), Abraham Becker, *Theorie der Elektrizitat*, vol ii, 6th ed , Leipzig, 1933

bodies, but instead of (72) Abraham assumed the electromagnetic momentum density to be given by

$$\mathbf{g} = \frac{1}{c}(\mathbf{E}\times\mathbf{H}) = \frac{\mathbf{S}}{c^2} \tag{81}$$

in the rest system. Since Abraham's tensor S_{ik}^{Abr} is symmetrical in S^0 it is symmetrical in any system S, but in any system other than S^0 the components of S_{ik}^{Abr} are not simply given by (69), (70), (71), and (81), the expression of S_{ik}^{Abr} in terms of the field variables containing also the velocity $\mathbf{u}$ of the matter in a complicated way. The four-force density f_i^{Abr} derived from this tensor by the equation

$$f_i^{\mathrm{Abr}} = -\frac{\partial S_{ik}^{\mathrm{Abr}}}{\partial x_k} \tag{82}$$

also deviates in general from (73) in a complicated way. In the rest system we obviously have

$$\mathbf{f}^{\mathrm{Abr}} = \mathbf{f}+\frac{\epsilon\mu-1}{c^2}\frac{\partial \mathbf{S}}{\partial t}, \qquad f_4^{\mathrm{Abr}} = f_4, \tag{83}$$

i.e. in the rest system Abraham's force density differs from Minkowski's expression by the term $\frac{\epsilon\mu-1}{c^2}\frac{\partial \mathbf{S}}{\partial t}$. This term is generally so small that an experimental verification would be very difficult.

Until quite recently, most physicists were inclined to adopt Abraham's theory. However, the question was never quite settled, and recently Tamm† has taken up the discussion again and he comes to the conclusion that Minkowski's expression for the energy-momentum tensor is correct.

In the first place it should be remarked that, since an electromagnetic field in ponderable matter is an essentially non-closed system, there is no *a priori* reason for the symmetry of the energy-momentum tensor. Abraham's main argument for a symmetrical tensor was that the quantities of the macroscopic theory must be derivable from the corresponding quantities in the electron theory by averaging over appropriate space-time regions and, since the microscopic energy-momentum tensor s_{ik} is symmetrical, the averaged tensor $\bar{s}_{ik}$ must also be symmetrical. But, as remarked by Tamm,‡ the macroscopic tensor S_{ik} is not simply the average of s_{ik}, S_{ik} must rather be defined so as to give the correct force density and the correct moment of force, i.e. we must have

$$f_i = -\frac{\partial S_{ik}}{\partial x_k} = -\frac{\partial \bar{s}_{ik}}{\partial x_k} \tag{84}$$

† Ig. Tamm, see ref., p. 204.
‡ Idem, private communication.

and, on account of (10),

$$d_{ik} = -\bar{x}_i \frac{\partial S_{kl}}{\partial x_l} + \bar{x}_k \frac{\partial S_{il}}{\partial x_l} + S_{ik} - S_{ki} = -\overline{x_i \frac{\partial s_{kl}}{\partial x_l}} + \overline{x_k \frac{\partial s_{il}}{\partial x_l}}. \tag{85}$$

From (84) we can only conclude that

$$S_{ik} = \overline{s_{ik}} + a_{ik},$$

where a_{ik} may possibly be a non-symmetrical tensor satisfying

$$\frac{\partial a_{ik}}{\partial x_k} = 0.$$

From (85) we see that S_{ik} will be symmetrical only if

$$-\overline{x_i \frac{\partial s_{kl}}{\partial x_k}} + \overline{x_k \frac{\partial s_{il}}{\partial x_l}} = -\bar{x}_i \frac{\partial \bar{s}_{kl}}{\partial x_l} + \bar{x}_k \frac{\partial \bar{s}_{il}}{\partial x_l} \tag{86}$$

and this is not necessarily the case

Further, Tamm could show that Abraham's expression in some special cases leads to wrong results, while Minkowski's expression for the energy-momentum tensor is in accordance with the electron theory. It must also be mentioned that Dallenbach† from perhaps not quite cogent arguments has given a general derivation of Minkowski's tensor from the electron theory In the following section we shall meet another strong argument in favour of Minkowski's theory

While thus the electromagnetic energy-momentum tensor is non-symmetrical, we may assume that the total energy-momentum tensor of matter and field is symmetric, since we then have to deal with a closed system. This means, however, that the mechanical energy-momentum tensor of the matter must also be non-symmetrical This is not in contradiction with the considerations in § 65, since we there considered a closed mechanical system, and the expression (VI 66) for the momentum density was derived from the explicit assumption $\mathbf{g} = \mathbf{S}/c^2$, which must be abandoned in our case if we adopt Minkowski's expression for the electromagnetic momentum density

76. The propagation velocity of the energy of a light wave in a moving refractive body

In Chapters I and II we have defined the direction and velocity of a light ray in a transparent refractive body by means of Huyghens's principle, and in § 24 it was shown that the ray velocity so defined transforms like the velocity of a particle by Lorentz transformations, i e by the equations (II 45–47) As a consequence of these equations we arrived

† See ref , p 195.

in § 25 at the aberration formula (II 91) and at Fresnel's formula (II 92) which, as regards effects of the first order, are in agreement with the experiments.

According to Maxwell's theory of light, optical phenomena in a refractive body with the refractive index n are described by means of Maxwell's phenomenological equations of electrodynamics for a substance with the electric and magnetic constants ϵ and μ connected with n by the equation

$$n = \sqrt{(\epsilon\mu)} \tag{87}$$

At least this is true for sufficiently long waves, where we can neglect all dispersion phenomena Further, a transparent body which does not absorb any light must be regarded as a perfect insulator, i e we have

$$\sigma = 0, \qquad \mathbf{J} = 0, \qquad \rho = 0 \tag{88}$$

Now the ray velocity must be identical with the velocity with which the energy in the wave is propagated In the aberration experiment, for instance, the angle of aberration is the angle through which the telescope must be tilted in order to get the ray, i e the energy, into the telescope The direction of the ray velocity must therefore be the same as the direction of propagation of the energy in the wave When the energy-momentum tensor of the electromagnetic field is given, we can, however, find the velocity of the energy by means of (VI 9), i e

$$\mathbf{u}^* = \mathbf{S}/h \tag{89}$$

We must therefore require that $\mathbf{u}^*$ in the case of a light wave transforms as a particle velocity by Lorentz transformations. This means that the quantity (VI 15) must be a four-vector As shown in § 62, this is the case only if the energy-momentum tensor satisfies the condition (VI 19). We shall now show that this condition is actually satisfied by Minkowski's tensor (68)-(72), but not by Abraham's tensor, and this is a strong argument in favour of Minkowski's theory

As mentioned in § 62, it is sufficient to prove the validity of the equation (VI 19) in one system We choose to work in the rest system of the refractive substance. We can obviously confine ourselves to the consideration of a plane wave, for in the problems which were considered in Chapters I and II the radii of curvature of the wave fronts are large compared with the wave-length (geometrical optics) and therefore the curved wave fronts can at each place be approximated by plane waves.

As shown in Appendix 3, the most general solution of the field equations with $\rho = 0$, $J = 0$, representing a plane wave with the wave normal **n** in the rest system, is

$$\left.\begin{aligned}\mathbf{E} &= \frac{f(t-(\mathbf{x}.\mathbf{n})/w)}{\sqrt{\epsilon}}\mathbf{e}^{(1)}+\frac{g(t-(\mathbf{x}.\mathbf{n})/w)}{\sqrt{\epsilon}}\mathbf{e}^{(2)}\\ \mathbf{H} &= -\frac{g(t-(\mathbf{x}.\mathbf{n})/w)}{\sqrt{\mu}}\mathbf{e}^{(1)}+\frac{f(t-(\mathbf{x}.\mathbf{n})/w)}{\sqrt{\mu}}\mathbf{e}^{(2)}\end{aligned}\right\}. \tag{90}$$

Here $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ are two fixed unit vectors which are perpendicular to each other and to the unit vector **n**, i.e.

$$(\mathbf{e}^{(1)}.\mathbf{e}^{(2)}) = (\mathbf{e}^{(1)}.\mathbf{n}) = (\mathbf{e}^{(2)}.\mathbf{n}) = 0, \qquad (\mathbf{e}_1\times\mathbf{e}_2) = \mathbf{n} \tag{91}$$

f and g are arbitrary functions of the argument $t-(\mathbf{x}.\mathbf{n})/w$, and

$$w = \frac{c}{\sqrt{(\epsilon\mu)}} = \frac{c}{n} \tag{92}$$

is the phase velocity.

Thus we get from Minkowski's expressions (68), (72) in the rest system

$$\mathbf{S} = c(\mathbf{E}\times\mathbf{H}) = \frac{c}{\sqrt{(\epsilon\mu)}}(f^2+g^2)\mathbf{e}, \tag{93}$$

where
$$\mathbf{e} = (\mathbf{e}_1\times\mathbf{e}_2) = \mathbf{n}, \tag{94}$$

i.e. the direction of propagation of the energy **e** coincides with the direction of the wave normal **n** in this system. Further, we have

$$h = \tfrac{1}{2}(\epsilon E^2+\mu H^2) = f^2+g^2 \tag{95}$$

Hence
$$\mathbf{u}^* = \frac{\mathbf{S}}{h} = \frac{c}{\sqrt{(\epsilon\mu)}}\mathbf{e} = u^*\mathbf{e} = \frac{c}{n}\mathbf{e} = \mathbf{w}, \tag{96}$$

i.e. the velocity of the energy is equal to the phase velocity in the rest system. From (96) we get, by means of (VI. 15 and 12),

$$U_\iota^* = \left(\frac{c}{\sqrt{(\epsilon\mu-1)}}\mathbf{e}, \frac{ic\sqrt{(\epsilon\mu)}}{\sqrt{(\epsilon\mu-1)}}\right), \tag{97}$$

$$S_\iota = (\mathbf{S}, ich) = (f^2+g^2)\left(\frac{c}{\sqrt{(\epsilon\mu)}}\mathbf{e}, ic\right) \tag{98}$$

Further, we get from (69)

$$\begin{aligned}S_{\iota\kappa} &= -t_{\iota\kappa} = -\epsilon E_\iota E_\kappa-\mu H_\iota H_\kappa+\tfrac{1}{2}(\epsilon E^2+\mu H^2)\delta_{\iota\kappa}\\ &= (f^2+g^2)(-e_\iota^{(1)}e_\kappa^{(1)}-e_\iota^{(2)}e_\kappa^{(2)}+\delta_{\iota\kappa}) = (f^2+g^2)e_\iota e_\kappa,\end{aligned} \tag{99}$$

since
$$e_\iota^{(1)}e_\kappa^{(1)}+e_\iota^{(2)}e_\kappa^{(2)}+e_\iota e_\kappa = \delta_{\iota\kappa}$$

on account of (91). Finally, (72) gives

$$S_{\iota 4} = icg_\iota, \qquad g_\iota = \frac{\epsilon\mu}{c}(\mathbf{E}\times\mathbf{H})_\iota = \frac{\sqrt{(\epsilon\mu)}}{c}(f^2+g^2)e_\iota. \tag{100}$$

For the quantity

$$a_\iota = \frac{S_{\iota l} U_l^*}{c^2} = \frac{1}{c^2}(S_{\iota\kappa} U_\kappa^* + icg_\iota U_4^*) \tag{101}$$

we then get, by means of (97), (99), and (100),

$$a_\iota = -\frac{(f^2+g^2)}{c}\sqrt{(\epsilon\mu-1)}e_\iota. \tag{102}$$

Thus the tensor R_{ik} defined in (VI. 19) has the following components:

$$\left.\begin{aligned} R_{\iota\kappa} &= S_{\iota\kappa}+a_\iota U_\kappa^* = (f^2+g^2)\left(e_\iota e_\kappa - \frac{\sqrt{(\epsilon\mu-1)}}{c}\cdot\frac{c}{\sqrt{(\epsilon\mu-1)}}e_\iota e_\kappa\right) = 0 \\ R_{\iota 4} &= icg_\iota + a_\iota U_4^* = i(f^2+g^2)\left(\sqrt{(\epsilon\mu)} - \sqrt{(\epsilon\mu-1)}\frac{\sqrt{(\epsilon\mu)}}{\sqrt{(\epsilon\mu-1)}}\right)e_\iota = 0 \\ R_{4k} &= 0 \end{aligned}\right\} \tag{103}$$

Hence the condition (VI. 19) is satisfied and the velocity of propagation of the energy $\mathbf{u}^*$ is in every system of coordinates identical with the ray velocity as determined by Huyghens's principle.†

If $\mathbf{v}'$ is the velocity of the rest system S relative to a system of coordinates S' with the same orientation of the spatial axes as in S, the transformation coefficients $\epsilon_{\iota\kappa}$ in (VI. 17′ or 18) are given by (IV. 126) with $\mathbf{v} = -\mathbf{v}'$, i.e.

$$\epsilon_{\iota\kappa} = 0, \qquad \epsilon_{\iota 4} = -\epsilon_{4\iota} = \frac{-iv'_\iota}{c}, \qquad \epsilon_{44} = 0 \tag{104}$$

Since the velocity of the energy in S' is

$$u_\iota^{*\prime} = \frac{S'_\iota}{h'} = \frac{icS'_\iota}{S'_4} = \frac{icU_\iota^{*\prime}}{U_4^{*\prime}}, \tag{105}$$

this quantity may be calculated from (VI. 18 or 17′), using (104). We then get

$$\mathbf{u}^{*\prime} = \frac{c}{\sqrt{(\epsilon\mu)}}\mathbf{e} + \mathbf{v}' - \frac{(\mathbf{v}'.\mathbf{e})\mathbf{e}}{\epsilon\mu}, \tag{106}$$

which, of course, is in accordance with (II. 55) if we put $\mathbf{v} = -\mathbf{v}'$ and $\mathbf{u} = \{c/\sqrt{(\epsilon\mu)}\}\mathbf{e}$ in this latter formula and neglect terms of order higher than the first in v'. For the magnitude of $\mathbf{u}^{*\prime}$ we get

$$u^{*\prime} = \frac{c}{n} + (\mathbf{e}.\mathbf{v}')\left(1 - \frac{1}{n^2}\right) \tag{107}$$

in accordance with the 'Fresnel formula' (II. 92).

† See also A. Scheye, *Ann. d. Phys.* (4), **30**, 805 (1909).

On the other hand, if we adopt Abraham's expressions for the energy-momentum tensor the equations (90)–(99) will still be valid in the rest system, but instead of (100) we get from (81)

$$\mathbf{g}^{\mathrm{Abr}} = \frac{1}{c}(\mathbf{E}\times\mathbf{H}) = \frac{\mathbf{S}}{c^2} = \frac{1}{c\sqrt{(\epsilon\mu)}}(f^2+g^2)\mathbf{e} = \mathbf{g}-(\epsilon\mu-1)\frac{\mathbf{S}}{c^2}. \quad (108)$$

Hence, for the quantity (101) we get in this case

$$a_i^{\mathrm{Abr}} = a_i - \frac{i}{c}(\epsilon\mu-1)\frac{S_i}{c^2}U_4^*,$$

and the tensor R_{ik} will not be zero any more Thus U_i^* will not be a four-vector and $\mathbf{u}^*$ will not transform like a particle velocity.

To find explicit expressions for the vector $\mathbf{u}^{*\prime}$ we again use (104), (105), and (VI 17′). Hence, with Abraham's expression for the energy-momentum tensor we get for the velocity of the electromagnetic energy in the system S'

$$\mathbf{u}^{*\prime} = \frac{c}{\sqrt{(\epsilon\mu)}}\mathbf{e}+\mathbf{v}' - \frac{(\mathbf{v}'.\mathbf{e})\mathbf{e}}{\epsilon\mu} + (\mathbf{v}'.\mathbf{e})\left(1-\frac{1}{\epsilon\mu}\right)\mathbf{e}, \quad (109)$$

which deviates from the ray velocity as defined by Huyghens's principle by the last term Instead of (107) we now have

$$u^{*\prime} = \frac{c}{n}+2(\mathbf{e}.\mathbf{v}')\left(1-\frac{1}{n^2}\right) \quad (109')$$

From (109) we see that the direction of the energy flux in this case is different from the direction of the ray velocity as defined by Huyghens's principle, which would give rise to a change in the aberration formula for light traversing a medium of refractive index $n > 1$, the deviation being of first order in v'. Unfortunately it is very difficult to measure even first-order aberration effects in moving transparent media

Further, we see from (109) that the velocity of the energy differs from the phase velocity even if $\mathbf{v}'$ and $\mathbf{e}$ are parallel, in contrast to the ray velocity which in this case is identical with the phase velocity This strange result is connected with the following circumstance

While Minkowski's four-force density f_i is zero in the case considered, Abraham's theory would give a non-vanishing force density on a homogeneous insulator In the rest system we have, according to (83),

$$\mathbf{f}^{\mathrm{Abr}} = \frac{\epsilon\mu-1}{c^2}\frac{\partial\mathbf{S}}{\partial t} = \frac{\epsilon\mu-1}{c\sqrt{(\epsilon\mu)}}\frac{\partial}{\partial t}(f^2+g^2)\mathbf{e}, \qquad f_4 = 0. \quad (110)$$

Thus, in this system, the electromagnetic energy is conserved, but this will not be the case in S'. From (104) and the transformation equations

$$f_i'^{\text{Abr}} = f_i^{\text{Abr}} + \epsilon_{ik} f_k^{\text{Abr}}$$

we get

$$f_4'^{\text{Abr}} = \epsilon_{4\kappa} f_\kappa^{\text{Abr}} = \frac{i}{c}\frac{\epsilon\mu-1}{c^2}\left(\mathbf{v}'\cdot\frac{\partial\mathbf{S}}{\partial t}\right) = \frac{i}{c}\frac{\epsilon\mu-1}{c\sqrt{(\epsilon\mu)}}(\mathbf{v}'\cdot\mathbf{e})\frac{\partial}{\partial t}(f^2+g^2), \tag{111}$$

which is i/c times the mechanical work on the substance per unit time and volume. Thus, in S' we have an exchange of energy between the electromagnetic and the mechanical system, i.e. a local absorption and re-emission of light energy by the body. This clearly shows that Minkowski's decomposition of the total energy-momentum tensor into an electromagnetic and a mechanical part is more natural than Abraham's, a transparent body being in Minkowski's theory a system which does not even locally exchange energy with the electromagnetic field.

77. The laws of thermodynamics in stationary matter

As shown by Planck† and Einstein‡, the usual laws of thermodynamics may be easily incorporated in the special theory of relativity. For simplicity we shall confine ourselves to the consideration of systems consisting of a thermodynamic fluid which can exert a normal pressure only on any surface element. In the rest system of the fluid the two laws of thermodynamics may be stated in the usual way.

According to the first law, the total energy of the system is a unique function of its state. In a thermodynamical process which gives rise to a change of state, the change in energy dE^0 is given by

$$dE^0 = \delta Q^0 + \delta A^0, \tag{112}$$

where δQ^0 is the amount of heat transferred to the system by the process, while δA^0 is the mechanical work done by the surroundings on the system. In a reversible process in which the volume V^0 of the system is increased by an infinitesimal amount, we have

$$\delta A^0 = -p^0\, dV^0, \tag{113}$$

where p^0 is the pressure.

According to the second law of thermodynamics, the entropy S^0 in the rest system is a function of the thermodynamical state. The change

† M. Planck, *Berl. Ber.*, p. 542 (1907), *Ann. d. Phys.* **76**, 1 (1908), F. Hasenohrl, *Wien Ber.* **116**, 1391 (1907).

‡ A. Einstein, *Jahrb. f. Rad. und El.* **4**, 411 (1907).

of entropy content of a system by a small change of state is defined by

$$dS^0 = \frac{\delta Q^0_{\mathrm{rev}}}{T^0} = \frac{dE^0 + p^0\,dV^0}{T^0}, \tag{114}$$

where $\delta Q^0_{\mathrm{rev}}$ is the amount of heat transferred to the system in a reversible process which brings about the change of state considered, and T^0 is the absolute temperature of the system. On the other hand, if the change of state is brought about by an irreversible process, we have always

$$dS^0 > \frac{\delta Q^0}{T^0}. \tag{115}$$

78. Transformation properties of the thermodynamical quantities

We shall now establish the transformation properties of the thermodynamical variables. Consider a system of inertia with respect to which our thermodynamical system is moving with the constant velocity $\mathbf{u}$. Since the fluid in question is in equilibrium in the rest system, the total momentum and energy are given by (VI 112, 113) which together with (VI 100) and (II 34) lead to the equations

$$\left.\begin{aligned} \mathbf{G} &= \frac{E^0+p^0V^0}{c^2\sqrt{(1-u^2/c^2)}}\mathbf{u}, & E &= \frac{E^0+p^0V^0\,u^2/c^2}{\sqrt{(1-u^2/c^2)}} \\ E+pV &= \frac{E^0+p^0V^0}{\sqrt{(1-u^2/c^2)}}, \quad p = p^0, & V &= V^0\sqrt{(1-u^2/c^2)} \end{aligned}\right\}. \tag{116}$$

Now consider again a reversible process which brings about a change in the state of the system. In an arbitrary system of coordinates the first law of thermodynamics may then again be written

$$dE = \delta Q + \delta A, \tag{117}$$

where δQ and δA are respectively the heat flow through the boundary and the work done by the surroundings on the system by the process considered. This work is, however, not simply equal to $-p\,dV$, as it was in the rest system. In order to preserve the equilibrium of the substance, i.e. a constant velocity throughout the body, we must assume the existence of external forces besides the normal pressure. The total extra force in the body, which is necessary to keep up a constant $\mathbf{u}$ in S, is

$$\mathbf{F} = \frac{d\mathbf{G}}{dt} = \frac{\mathbf{u}}{c^2\sqrt{(1-u^2/c^2)}}\frac{d}{dt}(E^0+p^0V^0), \tag{118}$$

and, since the right-hand side is different from zero during the thermodynamical process considered, there must be an external force on the system which carries out work of amount $(\mathbf{F}.\mathbf{u})\,dt = (\mathbf{u}\;d\mathbf{G})$ during the time dt. For the total work δA on the system we thus get

$$\delta A = -p\,dV + \mathbf{u}.d\mathbf{G}, \tag{119}$$

with $\mathbf{G}$ given by the first equation (116). From (117), (119), and (116) we thus get for constant $\mathbf{u}$

$$\begin{aligned}\delta Q &= dE + p\,dV - \mathbf{u}.d\mathbf{G} \\ &= \frac{dE^0 + (u^2/c^2)\,d(p^0V^0) - (u^2/c^2)(dE^0 + d(p^0V^0))}{\sqrt{(1-u^2/c^2)}} + p^0\,dV^0\sqrt{(1-u^2/c^2)} \\ &= (dE^0 + p^0\,dV^0)\sqrt{(1-u^2/c^2)}.\end{aligned} \tag{120}$$

Hence, by means of (112) and (113),

$$\delta Q = \delta Q^0\,\sqrt{(1-u^2/c^2)}. \tag{121}$$

This transformation formula for the amount of heat transferred through the boundary is identical with the formula (IV. 66) for the non-mechanical energy produced inside a body, for instance in the case of the Joule heating effect

Further, we define the entropy and absolute temperature in any system of coordinates by the equation

$$dS = \frac{\delta Q_{\text{rev}}}{T}. \tag{122}$$

Thus, for any reversible process without heat transfer, the entropy is constant. Consider now a thermodynamical system in some internal state originally at rest in a definite system of coordinates. If this system is accelerated reversibly and adiabatically to the velocity $\mathbf{u}$ without any change in its internal state, the entropy must be constant during the process on account of (122). Thus the entropy of the system is independent of the velocity when the internal state is the same, which means that the entropy must have the same value in every system of coordinates (cf. the analogous consideration, p. 141), i.e. S is an invariant

$$S = S^0. \tag{123}$$

From (122), (121), and (123) we then get the following transformation formula for the temperature

$$T = T^0\sqrt{(1-u^2/c^2)} \tag{124}$$

For any irreversible change of state we have in every system of coordinates

$$dS > \frac{\delta Q}{T}. \tag{125}$$

79. Four-dimensional formulation of the laws of thermodynamics

In § 69 it was shown that the laws of conservation of energy and momentum in differential form could be expressed by the tensor equation

$$\frac{\partial T_{ik}}{\partial x_k} = f_i. \tag{126}$$

This equation comprises the first law of thermodynamics. It is now possible also to express the second law of thermodynamics in four-dimensional language.†

Since the entropy is an additive quantity we can introduce the entropy density s defined so that $\delta S = s\,\delta V$ is the entropy content of the volume element δV. From (123) and the last equation (116) we get at once the transformation equation

$$s = \frac{s^0}{\sqrt{(1-u^2/c^2)}} \tag{127}$$

for the entropy density. For the change of entropy in the infinitesimal time δt of the element considered we then have, according to (125),

$$\frac{d}{dt}(s\,\delta V)\,\delta t \geqslant \frac{\delta Q}{T}, \tag{128}$$

where δQ is the heat transferred to the element in the time δt Here, $d(s\,\delta V)/dt$ is the substantial time derivative and, thus, we get by means of (IV. 200 and 204)

$$\frac{d}{dt}(s\,\delta V) = \frac{ds}{dt}\,\delta V + s\,\frac{d}{dt}\,\delta V = \left(\frac{\partial s}{\partial t} + (\mathbf{u}\,\mathrm{grad})s + s\,\mathrm{div}\,\mathbf{u}\right)\delta V$$

$$= \left(\mathrm{div}(s\mathbf{u}) + \frac{\partial s}{\partial t}\right)\delta V. \tag{129}$$

We can now introduce the four-current density of entropy

$$S_i = \left(\frac{s\mathbf{u}}{c}, is\right) \tag{130}$$

by analogy with the four-current density of electric charge (V 3) On account of (127) and (IV. 39) the four-vector S_i may also be written

$$S_i = \frac{s^0 U_i}{c} \tag{131}$$

by analogy with (V. 9).

† R C. Tolman, *Relativity, Thermodynamics and Cosmology*, § 71, Oxford 1934.

By means of (129) and (130) the left-hand side of (128) takes the form

$$c\,\frac{\partial S_\iota}{\partial x_\iota}\,\delta V\delta t = \frac{\partial S_\iota}{\partial x_\iota}\,\delta\Sigma, \tag{132}$$

where $\delta\Sigma = \delta V\delta x_4/i$ is the four-dimensional volume element defined by (IV. 121)

Since the right-hand side of (128) is an invariant on account of (121) and (124), the second law of thermodynamics may be written in the simple invariant form

$$\frac{\partial S_\iota}{\partial x_\iota}\,\delta\Sigma \geqslant \frac{\delta Q^0}{T^0}. \tag{133}$$

80. Ideal monatomic gases

For a monatomic gas consisting of N molecules we have in the rest system the usual expressions for the energy, entropy, etc., at least for moderate temperatures, where the mean kinetic energy of the molecules is small compared with the rest energy $m_0 c^2$ of a molecule. Thus if

$$\frac{kT^0}{m_0 c^2} \ll 1, \tag{134}$$

we have in the rest system the equation of state

$$p^0V^0 = NkT^0, \tag{135}$$

where k is Boltzmann's constant. Further,

$$E^0 = Nm_0c^2+\tfrac{3}{2}NkT^0, \tag{136}$$

where the energy has been normalized so that E^0 is equal to the sum of the rest energies of the molecules at zero temperature.

From (114), (135), and (136) we get in the usual way

$$dS^0 = \tfrac{3}{2}Nk\,\frac{dT^0}{T^0}+Nk\,\frac{dV^0}{V^0} \tag{137}$$

and by integration

$$S^0 = \tfrac{3}{2}Nk\ln T^0+Nk\ln V^0+C, \tag{138}$$

where C is a constant which is independent of T^0 and V^0.

In a system of coordinates in which the gas has the constant macroscopic velocity $\mathbf{u}$, we get by means of (116), (124), and (135)

$$pV = NkT, \tag{139}$$

i.e. the equation of state of an ideal gas is invariant in form. Further,

$$\left.\begin{aligned} \mathbf{G} &= \left(\frac{Nm_0}{\sqrt{(1-u^2/c^2)}}+\frac{\tfrac{5}{2}NkT}{c^2-u^2}\right)\mathbf{u} \\ E &= \frac{Nm_0c^2}{\sqrt{(1-u^2/c^2)}}+\frac{NkT(\tfrac{3}{2}+u^2/c^2)}{1-u^2/c^2} \\ S &= \tfrac{3}{2}Nk\ln T+Nk\ln V-\tfrac{5}{2}Nk\ln\sqrt{(1-u^2/c^2)}+C \end{aligned}\right\}. \tag{140}$$

The energy-momentum tensor is given by (VI. 104)

$$\left.\begin{aligned} T_{ik} &= \left(\mu^0+\frac{p^0}{c^2}\right)U_i U_k + p^0\,\delta_{ik}, \\ \text{with}\qquad \mu^0 &= \frac{E^0}{c^2V^0} = n^0\left(m_0+\frac{3}{2}\frac{kT^0}{c^2}\right), \qquad p^0 = n^0kT^0 \end{aligned}\right\}, \tag{141}$$

where $n^0 = N/V^0$ is the number of molecules per unit volume in the rest system. Thus we have the following relation between the pressure p^0, the rest density μ^0, and the temperature T^0

$$p^0 = \frac{\mu^0 kT^0}{m_0+\frac{3}{2}kT^0/c^2}. \tag{142}$$

81. Black-body radiation

The electromagnetic radiation inside a hollow enclosure in equilibrium with the walls at a definite temperature can be treated as a perfect fluid. In the rest system of the walls the flux of electromagnetic radiation is zero at every point and, according to Stefan–Boltzmann's law, the energy density h^0 is given by the formula

$$h^0 = aT^{0\,4}, \tag{143}$$

where

$$a = 7{\cdot}6237\times 10^{-15}\text{ erg cm.}^{-3}\text{ deg.}^{-4} \tag{144}$$

is Stefan–Boltzmann's constant. Thus h^0 is independent of the volume of the enclosure. The radiation exerts a normal pressure p^0 which is

$$p^0 = \frac{a}{3}T^{0\,4} = \frac{h^0}{3} = \frac{\mu^0c^2}{3}. \tag{145}$$

For the total energy of the radiation we now get

$$E^0 = aV^0T^{0\,4} \tag{146}$$

and, on account of (114),

$$dS^0 = 4aV^0T^{0\,2}\,dT^0+\tfrac{4}{3}aT^{0\,3}\,dV^0. \tag{147}$$

By integration we get

$$S^0 = \tfrac{4}{3}aV^0T^{0\,3}, \tag{148}$$

where the constant of integration has been chosen so as to make $S^0 = 0$ for $V^0 = 0$ and $T^0 = 0$.

In the system of coordinates, where the rest system is moving with velocity $\mathbf{u}$, we then get, on account of (116), (123), (124),

$$\left.\begin{aligned} \mathbf{G} &= \frac{4}{3}\frac{aV^0T^{0\,4}}{c^2\sqrt{(1-u^2/c^2)}}\mathbf{u} = \frac{4}{3}\frac{aVT^4}{c^2(1-u^2/c^2)^3}\mathbf{u} = \frac{4}{3}\frac{E^0/c^2}{\sqrt{(1-u^2/c^2)}}\mathbf{u} \\ E &= \frac{aV^0T^{0\,4}(1+\frac{1}{3}u^2/c^2)}{\sqrt{(1-u^2/c^2)}} = \frac{aVT^4(1+\frac{1}{3}u^2/c^2)}{(1-u^2/c^2)^3} = E_0\frac{1+\frac{1}{3}u^2/c^2}{\sqrt{(1-u^2/c^2)}} \\ S &= S^0 = \frac{4}{3}\frac{aVT^3}{(1-u^2/c^2)^2} \end{aligned}\right\}. \tag{149}$$

The relation between **G**, E, E^0, and **u** is the same as for a spherically symmetric electrostatic system (cf. (23)).

The energy-momentum tensor for black-body radiation is again

$$T_{ik} = \left(\mu^0 + \frac{p^0}{c^2}\right) U_i U_k + p^0 \delta_{ik} \tag{150}$$

with the relation (145) between the pressure and the density of the rest mass. It satisfies the same relation

$$T_{ii} = -\mu^0 c^2 + 3p^0 = 0 \tag{151}$$

as the electromagnetic energy-momentum tensor of an arbitrary electromagnetic field (cf. V. 108). This should also be expected, since the macroscopic energy-momentum tensor of the black-body radiation is the statistical average of the electromagnetic energy-momentum tensor in the canonical ensemble corresponding to the temperature T.

VIII

THE FOUNDATIONS OF THE GENERAL THEORY OF RELATIVITY

82. The general principle of relativity

According to the special principle of relativity which is the basis of the special theory of relativity, all systems of inertia, i.e. all rigid systems of reference moving with constant velocity relative to the fixed stars, are completely equivalent as regards our description of nature. Mathematically this principle found its expression in the covariance of the fundamental equations of physics under Lorentz transformations In spite of the inner consistency and harmony which characterize the special theory of relativity, it is, however, extremely unsatisfactory that this theory also distinguishes certain systems of reference, the systems of inertia, from all other conceivable systems of reference. This defect was felt especially serious in the treatment of the so-called clock paradox mentioned in § 20 There we actually had to refrain from giving a real solution of the clock paradox The question was simply rejected by a reference to the fact that the system of coordinates S^*, following the moving clock, is a system of inertia for only part of the time and that a discussion of the problem in this system of coordinates therefore falls beyond the scope of the special theory of relativity.

However, it seems to be rather difficult beforehand to acknowledge the accelerated systems of reference as equivalent to the systems of inertia as regards the description of natural phenomena (When, in the following chapters, we speak of an accelerated system, simply, we always have in mind a system accelerated relative to the systems of inertia or to the fixed stars) For example, if we consider a purely mechanical system consisting of a number of material particles acted upon by given forces, and with velocities small compared with the velocity of light relative to a system of inertia, Newton's fundamental equations of mechanics may be applied with good approximation in the description of this system. On the other hand, if we wish to describe the system in an accelerated system of reference, we must introduce, as is well known, so-called fictitious forces (centrifugal forces, Coriolis forces, etc.) which have no connexion whatever with the physical properties of the mechanical system itself. In fact, they depend exclusively on the acceleration of the system of reference introduced relative to the systems of inertia.

It was just for this reason that Newton introduced the concept of absolute space which should represent the system of reference where the laws of nature assume the simplest and most natural form. However, as mentioned at the beginning of Chapter II, the notion of absolute space lost its physical meaning as soon as the special principle of relativity was generally accepted, for as a consequence of this principle it became impossible by any experiment to decide which system of inertia had to be regarded as the absolute system Therefore, Einstein† advocated a new interpretation of the fictitious forces in accelerated systems of reference instead of regarding them as an expression of a difference in principle between the fundamental equations in uniformly moving and accelerated systems he considered both kinds of systems of reference to be completely equivalent as regards the form of the fundamental equations, and the 'fictitious' forces were treated as real forces on the same footing as any other force of nature. The reason for the occurrence in accelerated systems of reference of such peculiar forces should, according to this new idea, be sought in the circumstance that the distant masses of the fixed stars are accelerated relative to these systems of reference. The 'fictitious forces' are thus treated as a kind of gravitational force, the acceleration of the distant masses causing a 'field of gravitation' in the system of reference considered.

The idea that the acceleration of the distant masses can produce a gravitational field which is not perceptible in a system of inertia is not more artificial than, for example, the fact that an electrostatic system has zero magnetic field in the rest system of inertia of the charges, while a magnetic field is present in every system of inertia in which the charges are moving with constant velocity. The cause of the appearance of a magnetic field in the 'moving' system of inertia must be sought in the motion of the electric charges relative to these systems, and the appearance of the magnetic field can at any rate not be taken as an indication that the fundamental equations of electromagnetics have different forms in different systems of inertia The only essential difference between the two cases considered is the circumstance that the cause of the magnetic field can be found in the state of motion of terrestrial systems (viz. that of the charges), while the origin of the gravitational fields in accelerated systems must be sought in the state of motion of the distant celestial masses. Previously the effect of the celestial masses had been considered to be negligible; now, however, we must include

† A Einstein, *Jahrb f. Rad und El.* **4,** 411 (1907), *Ann d. Phys* **35,** 898 (1911), ibid. **38,** 355, 443 (1912), *Phys ZS* **14,** 1249 (1913)

the distant masses in the physical system considered. Only when we work in special systems of reference, viz. systems of inertia, is it not necessary to include the distant masses in our considerations, and this is the only point which distinguishes the systems of inertia from other systems of reference. *It can, however, be assumed that all systems of reference are equivalent with respect to the formulation of the fundamental laws of physics.* This is the so-called *general principle of relativity.*

83. The principle of equivalence

The interpretation of the 'fictitious forces' as gravitational forces is corroborated decisively by the fact that they have an essential property in common with the usual gravitational fields, viz. the property to give all free particles the same acceleration irrespective of the mass of the particles It is immediately clear that the 'fictitious forces' have this property, and Galileo was the first to prove this property for the gravitational field of the earth As a result of his experiments he was able to make the statement that *all bodies 'are falling with equal speed'* in empty space. This result simply expresses the fact that the force with which the gravitational field of the earth affects a particle is proportional to the inertial mass of the particle which determines the inertia of the particle against changes of motion As long as the velocity of the particle is small relative to the velocity of light, its motion in the direction of the gravitational field is therefore given by the equation

$$m\ddot{x} = mg,$$

where m is the mass of the particle and $\ddot{x}$ is the acceleration of the particle in the direction of the gravitational field The quantity g is a measure of the strength of the gravitational field and is independent of the mass of the particle This circumstance is frequently expressed by stating that *the ratio between the inertial mass of a particle and its gravitational mass is a universal constant*, depending only on the units in which the quantities in question are measured. This theorem has now been proved by a large number of experiments,† the most accurate of which are those performed by Eotvos and by Zeeman. The ratio of the inertial and gravitational mass was always found to be the same. A particular interest was attached to the experiments of Southerns and Zeeman with uranium, which at that time was known to have a great mass

† R v Eotvos, *Math u naturw Ber aus Ungarn*, **8**, 65 (1890), *Ann. d Phys.* **59**, 354 (1896). L. Southerns, *Proc Roy Soc London*, A, **84**, 325 (1910); P. Zeeman, *Proc. Amst* **20**, 542 (1917) R. v Eotvos, D. Pekár, and E. Fekete, *Ann. d Phys* **68**, 11 (1922).

defect. In Chapter III we have seen that any energy of amount E corresponds to an inertial mass $m = E/c^2$, a theorem which has been verified experimentally by numerous nuclear transformation processes (cf. § 32). The mass which is determined in a mass spectrograph obviously is the inertial mass, and Zeeman's result now shows that the binding energy of the uranium nucleus which appears in the mass defect also corresponds to a gravitational mass which has the same universal ratio to the inertial mass as for all other types of mass.

In view of the property just discussed, a gravitational field may thus be characterized by the 'gravitational acceleration' independent of the mass of the test particle, and this applies both to the usual gravitational fields due, for example, to the gravitation of the earth or the sun, and to those gravitational fields which appear in accelerated systems of reference and which are due to the distant masses of fixed stars Actually the gravitational field on the surface of the rotating earth is a mixture of these two types of field, the centrifugal force due to the rotation of the earth being in general not negligible compared with the force due to the attraction of the test body by the mass of the earth. It is thus quite natural to assume that both types of gravitational fields are of the same nature and obey the same fundamental laws. This assumption is often referred to as the *principle of equivalence*. It is true that the gravitational fields due to the distant masses can be made to disappear by a suitable choice of the system of reference, viz. by choosing a system of inertia as system of reference, while the gravitational fields arising from 'close' masses such as that of the earth or the sun cannot be 'transformed away' by a proper choice of the system of reference; the latter fields will therefore be referred to as *permanent gravitational fields*.

In this respect, however, the situation is quite similar to the case of the magnetic fields with which the gravitational fields were compared in § 82. In some cases, viz. when the charges producing the electromagnetic field have the same constant velocity relative to the fixed stars, it is possible completely to transform away the magnetic field by choosing the rest system of the charges as system of reference, for in this system the field will be a purely electrostatic field. In general it will, however, not be possible to choose a system of reference in which the magnetic field disappears everywhere in this system. Nevertheless, in this case the electromagnetic field is not considered as essentially different from the field of a system where the magnetic field can be transformed away. In all cases the electromagnetic field obeys the same fundamental equations, viz. Maxwell's equations.

The most important task will now be to look for the general fundamental laws which all types of gravitational fields must obey. In the first place, we must, however, try to find the field functions which can give an adequate description of the fields of gravitation. To this end we shall first consider the simple case where no permanent gravitational fields are present. By a suitable choice of the system of reference, viz. in a system of inertia, the field of gravitation will disappear and we can apply the laws of the special theory in this system A simple transformation to accelerated systems of reference permits the determination of the field quantities describing the gravitational fields in accelerated systems, and, according to the principle of equivalence, these quantities may be assumed to give a correct description also of the more general permanent fields of gravitation.

84. Uniformly rotating systems of coordinates. Space and time in the general theory of relativity

The development of the ideas underlying the general principle of relativity now leads, as we shall see, to a still more radical revision of our conceptions of space and time than that required by the special principle of relativity. To illustrate the character of the problems with which we are confronted, we shall start with the consideration of a very simple accelerated system of coordinates, viz. a rigid uniformly rotating system.

Let I denote a certain system of inertia in a part of space so far from all masses that all gravitational effects can be neglected; further, let X, Y, Z, T denote the usual space-time coordinates defined in the way discussed in §§ 16 and 17. Instead of Cartesian coordinates we can obviously just as well employ general curvilinear coordinates for the fixation of the points in physical space. If we confine ourselves to the consideration of events in the XY-plane we can, for instance, introduce polar coordinates (R, θ) by means of the equations

$$X = R\cos\theta, \qquad Y = R\sin\theta. \tag{1}$$

Now we can define a uniformly rotating system of coordinates S with spatial coordinates

$$x = r\cos\vartheta, \qquad y = r\sin\vartheta \tag{2}$$

by means of the transformation equations

$$r = R, \qquad \vartheta = \theta - \omega T. \tag{3}$$

Any fixed point in the rotating system corresponding to constant values of (x, y) or (r, ϑ) obviously performs a circular motion relative to I with the constant angular velocity ω. For $T = 0$ the two systems of coordinates coincide.

Thus for all points with

$$r = R < c/\omega \tag{4}$$

the rotating system of reference may be represented by a uniformly rotating material disk. Each point p on this disk is characterized by a pair of numbers (x, y) or (r, ϑ) which are equal to the coordinates (X, Y) or (R, θ) of that point in the fixed XY-plane with which the point p coincides at the moment when the clocks in the system of inertia I record the time $T = 0$.

For the measurement of distances between fixed points on the rotating disk we shall use standard measuring-rods of the same kind as those used in the systems of inertia, but now at rest relative to the rotating disk. In connexion with the process of measuring distances in accelerated systems of reference a problem arises which did not occur in inertial systems. If the measuring-rods in one way or other are kept in a fixed position relative to an accelerated system of reference, they will generally be submitted to forces which may cause a deformation of the measuring-rods; for, according to the special theory of relativity, no absolutely rigid bodies can exist, since they would provide a means of transmitting signals with velocities larger than c.

Consider, for example, a measuring-rod, one end of which is attached to the point (r, ϑ) on the rotating disk and which lies in the direction of the radius The centrifugal forces will then undoubtedly cause a lengthening of the measuring-rod This deformation will, however, depend on the elastic properties of the material from which the measuring-rods are made, and all such deformations of the measuring-rod can therefore easily be corrected for. Now we make the assumption that *measuring-rods on the disk,* after insertion of these corrections, *have exactly the same length relative to I as the standard measuring-rod in the system of inertia I^0 which* at the moment considered *has the same velocity as the measuring-rod on the rotating disk.* In general, we shall assume *that the (corrected) standard measuring-rods in an accelerated system relative to the measuring-rods in I are subjected to Lorentz contractions only, which means that the lengths of the rods are independent of the accelerations relative to I.*

If we therefore measure the distance between two points (r, ϑ) and $(r+dr, \vartheta)$ on the disk with a standard measuring-rod on the disk we get the value

$$d\sigma = dr; \tag{5}$$

for the velocity of the measuring-rod relative to I is perpendicular to the direction of the rod and thus does not give rise to any Lorentz contraction. On the other hand, if we consider two points with the co-ordinates (r, ϑ) and $(r, \vartheta+d\vartheta)$ on the disk, a measuring-rod connecting these two points will have the velocity $r\omega$ relative to I in its own direction and it will therefore be contracted relative to the measuring-rods in I, in accordance with Lorentz's formula (II. 33). Hence the distance between the two points measured with the contracted measuring-rod will be

$$d\sigma = \frac{r\,d\vartheta}{\sqrt{(1-r^2\omega^2/c^2)}}. \tag{6}$$

For the distance $d\sigma$ between two neighbouring points (r, ϑ) and $(r+dr, \vartheta+d\vartheta)$, measured with a standard measuring-rod on the disk, we get similarly

$$d\sigma^2 = dr^2 + \frac{r^2\,d\vartheta^2}{1-r^2\omega^2/c^2}. \tag{7}$$

This follows immediately by means of the Pythagorean theorem if we consider the measuring procedure from the point of view of an observer in the system of inertia I, taking into account the fact that the measuring-rod moving with the rotating disk relative to the measuring-rods in I is contracted in the direction of the velocity, but unchanged in a direction perpendicular to the direction of motion.

It is now immediately clear that the geometrical theorems obtained by means of measurements with standard measuring-rods at rest relative to the disk in general will deviate from the theorems of Euclidian geometry. Consider, for example, the curve given by the equation

$$r = \text{constant}.$$

According to (5) this curve will represent a circle with radius r. The periphery of this circle will, however, according to (6), have the length

$$\int_0^{2\pi} \frac{r\,d\vartheta}{\sqrt{(1-r^2\omega^2/c^2)}} = \frac{2\pi r}{\sqrt{(1-r^2\omega^2/c^2)}}. \tag{8}$$

Therefore the ratio of the length of the periphery to the radius will not be 2π, but

$$\frac{2\pi}{\sqrt{(1-r^2\omega^2/c^2)}} > 2\pi.$$

Consequently we see that the recognition of the general principle of relativity, according to which the accelerated systems of coordinates are equivalent to the systems of inertia for the description of nature, forces us in some cases to abandon the Euclidean geometry which, as particularly advocated by Kant and even in the special theory of relativity, was regarded as an indispensable foundation of all description of space. This also has the consequence (cf. § 87) that it is not possible in general to use Cartesian space coordinates in accelerated systems of reference and that we have to make use of general curvilinear coordinates for the specification of the points in physical space.

Similarly, the general principle of relativity also requires a renewed revision of the notion of time. In the special theory of relativity the time in a system of inertia was simply defined by means of standard clocks placed at different points in the system and regulated by means of light signals in the way described in § 16. Two clocks which had been synchronized in this way remained synchronous. Likewise we could now think of defining the time in the rotating system of coordinates by means of standard clocks inserted at rest everywhere in this system and set according to the standard clocks in the system of inertia I, for example, by putting them to zero at the moments when the clocks in I with which they coincide show zero. Then we have for the time t thus defined in the rotating system $t = 0$ for $T = 0$. A standard clock at the point r, ϑ on the disk has, however, a velocity $r\omega$ relative to I, and therefore it will be retarded relative to the clocks in I in agreement with equation (II. 36), i.e. at a later time we have

$$t = T\sqrt{(1-r^2\omega^2/c^2)}. \tag{9}$$

In accordance with the assumption made in § 20, it is implied that only the velocity, not the accelerations relative to I, will influence the rate of a standard clock. Actually, sufficiently strong accelerations will of course more or less influence the rate of a real clock (cf. a watch which is dropped on the floor), but such an effect which depends on the material of which the clock is made can be corrected for, just as was the case with the measuring-rods.

The description of time in the rotating system which we obtain by using the time variable t defined by (9) is, however, although admissible in principle, highly unpractical. Imagine, for instance, a light source (an atom) which is placed at the point A with coordinates (r, ϑ) and which emits light with the proper frequency ν_0. The number of waves emitted in the time-interval from $t = 0$ to $t = 1$ is then by definition equal to ν_0. The number of waves emitted during the time from $T = 0$ to $T = 1$ is therefore, according to (9), $\nu_0(1-r^2\omega^2/c^2)^{\frac{1}{2}}$. The same number of waves will also arrive at the centre 0 ($r = 0$) during the time-interval from $T = 0$ to $T = 1$ or, since $t = T$ for $r = 0$, in the time-interval from $t = 0$ to $t = 1$.

The number of light waves which are emitted from the point A per unit time in the time-scale t is thus larger than the number of waves arriving at the centre 0 per unit time, and with this time variable we get a very complicated description of the propagation of light. This consideration shows that in general it is not convenient in accelerated systems of reference to use a time variable defined by standard clocks, and that a much simpler description may be obtained when one uses clocks of a different rate. In the case of the rotating disk, for example, it will be most convenient to use *coordinate clocks* whose rate at any place is $(1-r^2\omega^2/c^2)^{-\frac{1}{2}}$ times faster than the rate of the corresponding standard clock, for this means that the time parameter t defined by these coordinate clocks is identical with the time T in I, i.e. we have the transformation

$$t = T \tag{10}$$

instead of (9).

In principle it is admissible, however, to use coordinate clocks of an arbitrary rate, provided that the time variable t defined by these coordinate clocks gives a reasonable chronological ordering of the physical events. In accelerated systems of reference the spatial and temporal coordinates thus lose every physical significance, they simply represent a certain arbitrary, but unambiguous, numbering of the physical events.

85. Non-Euclidean geometry. The metric tensor

As we have seen, the spatial geometry on the rotating disk is non-Euclidean. Although all geometrical experience in the three-dimensional physical space is in complete agreement with the theorems of Euclidean geometry, the notion of non-Euclidean geometry in two dimensions is by no means foreign to us, since we meet examples of such geometries on every curved surface A well-known example is the spherical geometry on spherical surfaces. As an introduction to the non-Euclidean geometries in n-dimensional space, we shall therefore consider the geometry on an arbitrary two-dimensional surface embedded in a three-dimensional Euclidean space. If x, y, z are Cartesian coordinates in this space, the surface in question may be given by a parametric representation

$$x = F(x^1,x^2), \qquad y = G(x^1,x^2), \qquad z = H(x^1,x^2), \tag{11}$$

where F, G, H are given functions of the two parameters x^1 and x^2 in certain intervals. By differentiation of (11) we obtain

$$\left.\begin{aligned} dx &= \frac{\partial F}{\partial x^1}\,dx^1+\frac{\partial F}{\partial x^2}\,dx^2 \\ dy &= \frac{\partial G}{\partial x^1}\,dx^1+\frac{\partial G}{\partial x^2}\,dx^2 \\ dz &= \frac{\partial H}{\partial x^1}\,dx^1+\frac{\partial H}{\partial x^2}\,dx^2 \end{aligned}\right\}. \tag{11'}$$

The distance ds between two adjacent points on the surface corresponding to the parameter values (x^1, x^2) and (x^1+dx^1, x^2+dx^2), respectively, is given by

$$ds^2 = dx^2+dy^2+dz^2,$$

where dx, dy, dz are linear expressions in dx^1 and dx^2 given by (11′). Using these expressions we therefore get ds^2 expressed as a homogeneous quadratic form in dx^1 and dx^2, i.e.

$$ds^2 = g_{11}(dx^1)^2+g_{12}(dx^1dx^2)+g_{21}(dx^2dx^1)+g_{22}(dx^2)^2 \tag{12}$$

with

$$\left.\begin{aligned} g_{11} &= \left(\frac{\partial F}{\partial x^1}\right)^2+\left(\frac{\partial G}{\partial x^1}\right)^2+\left(\frac{\partial H}{\partial x^1}\right)^2 \\ g_{12} = g_{21} &= \frac{\partial F}{\partial x^1}\frac{\partial F}{\partial x^2}+\frac{\partial G}{\partial x^1}\frac{\partial G}{\partial x^2}+\frac{\partial H}{\partial x^1}\frac{\partial H}{\partial x^2} \\ g_{22} &= \left(\frac{\partial F}{\partial x^2}\right)^2+\left(\frac{\partial G}{\partial x^2}\right)^2+\left(\frac{\partial H}{\partial x^2}\right)^2 \end{aligned}\right\}. \tag{12′}$$

The angle θ between two line elements corresponding to the increments $(\delta x^\iota) = (\delta x^1, \delta x^2)$ and $(\Delta x^\iota) = (\Delta x^1, \Delta x^2)$, respectively, for the parameters (x^1, x^2) is given by the equation

$$\cos\theta = \frac{\delta x\Delta x+\delta y\Delta y+\delta z\Delta z}{\delta s\Delta s}, \tag{13}$$

where $(\delta x, \delta y, \delta z)$ and $(\Delta x, \Delta y, \Delta z)$ are the increments in x, y, z obtained from (11′) by replacing $dx^\iota = (dx^1, dx^2)$ by (δx^ι) and (Δx^ι), respectively, δs and Δs being the lengths of the line elements. Applying (11′), (12), and (12′) this equation can be written in the form

$$\cos\theta = \frac{g_{11}\,\delta x^1\Delta x^1+g_{12}\,\delta x^1\Delta x^2+g_{21}\,\delta x^2\Delta x^1+g_{22}\,\delta x^2\Delta x^2}{\delta s\Delta s}. \tag{14}$$

The line elements $(\delta x^\iota) = (\delta x^1, \delta x^2)$ and $(\Delta x^\iota) = (\Delta x^1, \Delta x^2)$ also define an infinitesimal parallelogram on the surface with the area

$$d\sigma = \delta s\Delta s\sin\theta, \tag{14′}$$

where θ is given by (14).

The curves on the surface which are obtained from (11) by putting

$$x^1 = \text{constant}, \tag{15a}$$

$$x^2 = \text{constant}, \tag{15b}$$

respectively, are called coordinate curves. Every point on the surface is the point of intersection of two coordinate curves from the manifolds (15a) and (15b), respectively. If the line elements (δx^i) and (Δx^i) are lying in the directions of these coordinate curves we have

$$(\delta x^i) = (dx^1, 0) \quad \text{and} \quad (\Delta x^i) = (0, dx^2),$$

and we get from (12) and (14)

$$\delta s = (g_{11})^{\frac{1}{2}}\,dx^1, \qquad \Delta s = (g_{22})^{\frac{1}{2}}\,dx^2,$$

$$\cos\theta = \frac{g_{12}}{(g_{11}\,g_{22})^{\frac{1}{2}}}, \qquad \sin\theta = (1-\cos^2\theta)^{\frac{1}{2}} = (g/g_{11}\,g_{22})^{\frac{1}{2}},$$

where
$$g = g_{11}\,g_{22}-g_{12}^2 = \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} = |g_{ik}| \tag{16}$$

is the determinant corresponding to the scheme of numbers g_{ik}. For the area $d\sigma$ of the parallelogram we thus get from (14′)

$$d\sigma = \sqrt{g}\,dx^1dx^2. \tag{17}$$

If the parametric representation (11) of the surface has the property that every set of values of the parameters $x^i = (x^1, x^2)$ corresponds to one and only one point on the surface, $x^i = (x^1, x^2)$ represents a set of general curvilinear coordinates (Gaussian coordinates) on the two-dimensional surface. All the fundamental geometrical quantities can then be expressed in terms of these coordinates alone without reference to the variables of the three-dimensional space in which the surface was supposed to be embedded. If $g_{ik} = g_{ik}(x^l)$ are given functions of the coordinates (x^l), the line element is given by (12),

$$ds^2 = g_{ik}\,dx^i dx^k, \tag{18}$$

summation over i and k for the values 1 and 2 being implied in this expression.

The angle θ between the two line elements (δx^i) and (Δx^i) is given by (14),

$$\cos\theta = \frac{g_{ik}\,\delta x^i\Delta x^k}{\sqrt{(g_{ik}\,\delta x^i\delta x^k)}\sqrt{(g_{ik}\,\Delta x^i\Delta x^k)}}, \tag{19}$$

and the area of the parallelogram defined by two line elements in the directions of the coordinate curves is given by (17). The quantities g_{ik}—the components of the so-called *metric tensor*—thus determine the geometry on the surface which in general will be non-Euclidean.

86. Geodesic lines

The straight lines which in Euclidean geometry may be defined as the curves of shortest distance are in the more general case replaced by the geodesic lines which likewise may be defined by a variational principle.

Let us consider an arbitrary curve connecting two points P_1 and P_2 on the two-dimensional surface. In a parametric representation of this curve the Gaussian coordinates $x^i = (x^1, x^2)$ may be regarded as certain functions of an arbitrary parameter λ in the interval

$$\lambda_1 < \lambda < \lambda_2,$$

i.e.

$$x^i = x^i(\lambda) \qquad (i = 1, 2). \tag{20}$$

Let $L = L(x^i, \dot{x}^i)$ be a given function of the variables

$$x^i \quad \text{and} \quad \dot{x}^i = dx^i/d\lambda.$$

The curve (20) which gives the integral

$$\int_{\lambda_1}^{\lambda_2} L(x^i, \dot{x}^i)\, d\lambda$$

a stationary value for all infinitesimal variations of the curve connecting the fixed points P_1 and P_2 is then determined by the condition

$$\delta \int_{\lambda_1}^{\lambda_2} L(x^i, \dot{x}^i)\, d\lambda = 0 \tag{21}$$

for all variations $\delta x^i(\lambda)$ satisfying the boundary condition

$$\delta x^i(\lambda_1) = \delta x^i(\lambda_2) = 0. \tag{22}$$

Now we have

$$\delta \int_{\lambda_1}^{\lambda_2} L\, d\lambda = \int_{\lambda_1}^{\lambda_2} \left\{ \frac{\partial L}{\partial x^i}\, \delta x^i(\lambda) + \frac{\partial L}{\partial \dot{x}^i}\, \delta \dot{x}^i(\lambda) \right\} d\lambda, \tag{23}$$

and since $\delta \dot{x}^i = d(\delta x^i)/d\lambda$, we obtain by partial integration of the last term in (23), taking into account the boundary condition (22),

$$\delta \int_{\lambda_1}^{\lambda_2} L\, d\lambda = \int_{\lambda_1}^{\lambda_2} \left[\frac{\partial L}{\partial x^i} - \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^i}\right) \right] \delta x^i(\lambda)\, d\lambda.$$

This integral can only be zero for all imaginable variations $\delta x^i(\lambda)$ if the factor in the square brackets is zero along the whole curve. The variational principle (21), (22) is thus equivalent to the *Euler differential equations*

$$\frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i} = 0, \qquad \dot{x}^i = \frac{dx^i}{d\lambda}. \tag{24}$$

If L is a homogeneous function of the variables $\dot{x}^i$ of nth degree, we have

$$\dot{x}^i \frac{\partial L}{\partial \dot{x}^i} = nL, \tag{25}$$

and for $n \neq 1$ it is easily seen that the function $L(x^i, \dot{x}^i)$ must be constant along the curve defined by the differential equations (24). From (24) and (25) we get

$$\frac{dL}{d\lambda} = \frac{\partial L}{\partial x^i}\dot{x}^i + \frac{\partial L}{\partial \dot{x}^i}\ddot{x}^i = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\dot{x}^i + \frac{\partial L}{\partial \dot{x}^i}\frac{d\dot{x}^i}{d\lambda} = \frac{d}{d\lambda}\left(\frac{\partial L}{\partial \dot{x}^i}\dot{x}^i\right) = n\frac{dL}{d\lambda}. \tag{26}$$

Hence $$(n-1)\frac{dL}{d\lambda} = 0. \tag{27}$$

If $n \neq 1$, we can thus conclude that

$$L(x^i, \dot{x}^i) = \text{constant} \tag{28}$$

is an integral of the equations (24).

Now we define the geodesic lines by the variational principle (21) with

$$L(x^i, \dot{x}^i) = g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} = \left(\frac{ds}{d\lambda}\right)^2. \tag{29}$$

The corresponding Euler equations (24) are

$$\frac{d}{d\lambda}\left(g_{ik}\frac{dx^k}{d\lambda}\right) = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda}, \tag{30}$$

which represent two differential equations of second order for the two functions $x^i(\lambda)$. Since L in (29) is homogeneous of the second degree in $\dot{x}^i$,

$$L = g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} = \left(\frac{ds}{d\lambda}\right)^2 = \text{constant} \tag{31}$$

is an integral of (30), and by a suitable choice of the parameter λ we can always choose the constant on the right-hand side of (31) to be equal to 1. This obviously means that we choose the length s of the curve measured along the geodesic line itself as parameter. (30) and (31) then take the form

$$\frac{d}{ds}\left(g_{ik}\frac{dx^k}{ds}\right) = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\frac{dx^k}{ds}\frac{dx^l}{ds}, \qquad g_{ik}\frac{dx^i}{ds}\frac{dx^k}{ds} = 1. \tag{32}$$

It is now seen at once that the curves defined by (30) and (31) also satisfy the Euler equations (24) with

$$L = \left(g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda}\right)^{\frac{1}{2}},$$

which means that the geodesic lines also satisfy the variational equation

$$\delta\int\left(g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda}\right)^{\frac{1}{2}}d\lambda = \delta\int ds = 0, \tag{33}$$

i.e. the distance between two points measured along the geodesic line connecting two arbitrary points has a stationary value.

Geodesic lines, angles between two intersecting geodesic lines as well as the distance between two points measured along the connecting geodesic line, are now completely determined by (32), (19), and (18) when the metric tensor g_{ik} is given as a function of the general coordinates (x^i). Then we can also deduce geometrical theorems regarding triangles formed by geodesic lines, etc., on the surface, i.e. the geometry of the two-dimensional space is completely defined by the quantities $g_{ik}(x^l)$.

87. Determination of the metric tensor by direct measurements. Geometry in n-dimensional space

The preceding deductions are, of course, completely independent of the system of coordinates employed If we use another set of curvilinear coordinates x'^i connected with the original coordinates x^i by a transformation

$$x'^i = f^i(x^1, x^2), \tag{34}$$

we have

$$dx'^i = \frac{\partial f^i}{\partial x^k} dx^k, \tag{35}$$

i.e. the differentials of the coordinates are connected by linear equations. Therefore if we eliminate the dx^i by means of (35) in (18), ds^2 will also be a homogeneous quadratic form in dx'^i

$$ds^2 = g_{ik}\, dx^i dx^k = g'_{ik}\, dx'^i dx'^k, \tag{36}$$

where the coefficients g'_{ik} can be regarded as functions of the new coordinates x'^i

Since the geodesic lines are defined by the invariant variational principle (21), (29), it is obvious that the differential equations for the geodesic lines expressed in the new coordinates are obtained from (30) or (32) by simply replacing g_{ik} and x^i by g'_{ik} and x'^i, respectively, in other words, the equations (30) are covariant or form-invariant. The same holds for the equation (19).

If it is possible by a transformation of the type (34) to introduce coordinates $X^i = f^i(x^k)$ such that the line element in the new coordinates assumes the form

$$ds^2 = (dX^1)^2+(dX^2)^2 = \delta_{ik}\, dX^i dX^k, \tag{37}$$

the geometry on the surface is called Euclidean. In this case the coordinates X^i play the same part as do Cartesian coordinates in a Euclidean plane. The differential equations (32) for the geodesic lines reduce in these coordinates to the equations

$$\frac{d^2X^i}{ds^2} = 0, \tag{38}$$

which have the same form as the equations for straight lines in Cartesian coordinates. Examples of such surfaces are cylinders and cones which can be unfolded on a plane without internal deformation. All geometrical theorems about triangles or other figures on such surfaces are identical with the theorems of Euclidean geometry, and if we are interested only in the two-dimensional geometry on the surface, all such surfaces may be regarded as identical.

In general it is not possible to introduce on the surface such coordinates (Cartesian coordinates) for which the line element assumes the form (37). In this case the geometry on the surface is a non-Euclidean general Riemannian geometry. In any case it is possible by means of measurements on the surface to determine the geometry on the surface without referring to the three-dimensional Euclidean space in which the surface is embedded. Let us assume that we have introduced an arbitrary system of coordinates, x^i, i e. an arbitrary continuous one-to-one correspondence of the set of numbers (x^i) and the points on the surface. By means of a measuring-rod we can now measure the distance ds between the points x^i and x^i+dx^i. Since then ds and dx^i are known numbers, the equation (18) represents one equation for the determination of the unknowns g_{ik}. Since the metric tensor in two dimensions has three independent components we may thus—by performing this procedure for three suitable line elements starting from the same point x^i—completely determine the values of g_{ik} at this point. This can now be done for any point on the surface, thus obtaining a complete experimental determination of the metric tensor

In this way the geometry on the surface becomes an empirical science subjected to the limitations arising from the limited measuring accuracy Now imagine that we heat the surroundings of a given point on the surface so that the measuring-rods are dilated when inserted at this point. If we neglect this dilatation we shall, by means of the method described above, find wrong values for the components of the metric tensors. Since, however, the thermal dilatation is different for measuring-rods prepared from various materials, it is easy to correct for this error and to find the 'real' values for the components of the metric tensor. On the other hand, it is obvious that, if all measuring-rods in the neighbourhood of a given point for one or another reason were dilated at the same rate independently of the material of which they are made, it would be impossible to observe this dilatation and in an unambiguous way to correct for it. Therefore there is no well-defined meaning in the statement that such an expansion of the standard rods has taken place, and from a physicist's point of view the metric tensor and the geometry obtained by measurements with the natural standard rods must be the 'real' geometry on the surface.

All the considerations of this paragraph for the two-dimensional case can now be immediately generalized to spaces of 3, 4, or n dimensions. The only difference is that the points in an n-dimensional space are characterized by n coordinates x^i and that all indices in the preceding equations now can take on the values from 1 to n. The length of a line element and the angle between two line elements are then again given by (18) and (19), respectively, and the geodesic lines are defined by the n equations (30) or by means of the variational principle expressed by the equations (21), (22), (29).

88. General accelerated systems of reference. The most general admissible space-time transformations

In § 84 we have seen that the spatial geometry in a uniformly rotating system of reference is non-Euclidean and also the temporal description is more complicated than in the systems of inertia. This may be regarded as an effect of the gravitational field present in the rotating system of reference According to the principle of equivalence we must therefore expect that a gravitational field in general will manifest itself not only by the presence of gravitational forces (centrifugal forces, Coriolis forces, gravitational attraction between masses, etc), but also in the results of space and time measurements.

Let us start again from a system of inertia I with the usual space and time coordinates (X, Y, Z, T) To each set of values of these variables corresponds a certain event which is represented by a point in (3+1)-space with coordinates

$$X^i = (X, Y, Z, cT). \tag{39}$$

These coordinates differ from the coordinates defined in (IV. 2) only in that we have dropped the i in the fourth coordinate. The four-dimensional line element (IV. 26) thus takes the form

$$ds^2 = dX^2+dY^2+dZ^2-c^2\,dT^2 = G_{ik}\,dX^i dX^k, \tag{40}$$

where

$$G_{ik} = \begin{cases} 0 \text{ for } i \neq k \\ 1 \text{ for } i = k = 1, 2, 3 \\ -1 \text{ for } i = k = 4. \end{cases} \tag{41}$$

Instead of the pseudo-Cartesian coordinates X^i we shall now introduce general 'curvilinear' coordinates x^i in four-space by means of the transformations

$$x^i = x^i(X^k), \tag{42}$$

where the $x^i(X^k)$ are arbitrary continuous and differentiable functions of the variables (X^k). The transformations (1), (3), (10) obviously represent a special case of (42). By differentiation of (42) we obtain

$$dx^i = \frac{\partial x^i}{\partial X^k} dX^k = \mathrm{A}^i_k\, dX^k. \tag{43}$$

From the relations reciprocal to (42) we obtain in the same way

$$dX^i = \frac{\partial X^i}{\partial x^k} dx^k = \breve{\mathrm{A}}^i_k\, dx^k. \tag{44}$$

Using (44) in the right-hand side of (43) gives

$$dx^i = \mathrm{A}^i_l \breve{\mathrm{A}}^l_k\, dx^k.$$

Since this equation is to hold for arbitrary dx^i, we must have

$$\mathrm{A}^i_l \breve{\mathrm{A}}^l_k = \delta^i_k = \begin{cases} 0 \text{ for } i \neq k \\ 1 \text{ for } i = k. \end{cases} \tag{45}$$

In the same way, substituting from (43) in (44), we obtain the equations

$$\breve{\mathrm{A}}^i_l \mathrm{A}^l_k = \delta^i_k \tag{45'}$$

Elimination of dX^i in (40) thus gives the following expression for the interval:

$$ds^2 = g_{ik}\, dx^i dx^k, \tag{46}$$

$$g_{ik} = g_{ki} = G_{lm} \breve{\mathrm{A}}^l_i \breve{\mathrm{A}}^m_k = \breve{\mathrm{A}}^\iota_i \breve{\mathrm{A}}^\iota_k - \breve{\mathrm{A}}^4_i \breve{\mathrm{A}}^4_k \tag{47}$$

(Greek indices run from 1 to 3, Latin indices from 1 to 4).

The system of coordinates (x^i) determined by the transformations (42) also determines a definite system of reference. Defining a 'point of reference' as a point with constant values for the three space-coordinates (x^ι), the system of reference corresponding to the system of coordinates (x^i) can be defined as the collection of all points of reference. In general such a system of reference will, of course, not be rigid, since the different reference points may have largely varying velocities relative to I. The motion of the different points in the system of reference will therefore in general be analogous to the motion of a fluid and we shall confine ourselves to such transformations (42) for which the corresponding system of reference can be pictured by a real fluid. This means that the velocities of the points of reference relative to the system of inertia I must always be smaller than c. Since for any point of reference $dx^\iota = 0$ we obtain from (44) for the velocity components v^ι of such a point relative to I,

$$\frac{v^\iota}{c} = \frac{dX^\iota}{dX^4} = \frac{\breve{\mathrm{A}}^\iota_4}{\breve{\mathrm{A}}^4_4}. \tag{48}$$

If this velocity is to be smaller than c, we must have

$$\frac{v^2}{c^2} = \frac{v^\iota v^\iota}{c^2} < 1 \quad \text{or} \quad \bar{A}^\iota_4 \bar{A}^\iota_4 - \bar{A}^4_4 \bar{A}^4_4 < 0. \tag{49}$$

In order that the system of reference defined by (42) may be physically realizable, the admissible space-time transformations must satisfy the condition (49) which, on account of (47), involves

$$g_{44} < 0. \tag{50}$$

At every point of reference we imagine a *coordinate* clock to be inserted showing the time $t = x^4/c$, and we must now further demand that the time description obtained in this way shall give a reasonable causal description of physical phenomena. Thus a real signal which is emitted from a point of reference (x^ι) at the time t must arrive at a point of reference $(x^\iota + dx^\iota)$ at a time $t+dt$ with positive dt. Since signals can at most have the velocity c relative to I, the line element

$$ds^2 = dX^\iota dX^\iota - c^2\, dT^2$$

must be smaller than or equal to zero for two adjacent points on the time track of the signal. In the system of coordinates x^ι this means

$$ds^2 = g_{ik}\, dx^i dx^k \leqslant 0.$$

In other words, any two events which are simultaneous in the system of coordinates (x^i), i.e. for which $dx^4 = 0$, cannot be connected by a signal, so that in this case we must have

$$g_{ik}\, dx^i dx^k > 0,$$

or, since $dx^4 = 0$,

$$g_{\iota\kappa}\, dx^\iota dx^\kappa > 0. \tag{51}$$

This inequality must now hold for arbitrary dx^ι, which shows that the quadratic form $g_{\iota\kappa}\, dx^\iota dx^\kappa$ must be *positive definite*. The necessary and sufficient condition for this to be true is that all subdeterminants of the scheme of numbers $g_{\iota\kappa}$ are positive.

The admissible transformations (42) must therefore be such that the corresponding g_{ik} satisfy the conditions

$$g_{\iota\iota} > 0, \qquad \begin{vmatrix} g_{\iota\iota} & g_{\iota\kappa} \\ g_{\kappa\iota} & g_{\kappa\kappa} \end{vmatrix} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, \qquad g_{44} < 0, \tag{52}$$

where ι and κ may be any of the numbers 1, 2, 3 (no summation over ι in $g_{\iota\iota}$).

From (52) it follows that the determinant

$$g = |g_{ik}| = \begin{vmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{vmatrix} < 0. \tag{53}$$

If we introduce another system of coordinates x'^i by means of the transformation equations

$$\left.\begin{aligned} x'^i &= x'^i(x^k) \\ dx'^i &= \alpha^i_k\, dx^k = \frac{\partial x'^i}{\partial x^k} dx^k \\ dx^i &= \check{\alpha}^i_k\, dx'^k = \frac{\partial x^i}{\partial x'^k} dx'^k \end{aligned}\right\}, \tag{54}$$

we have, by analogy with (45), (45′),

$$\alpha^i_l \check{\alpha}^l_k = \check{\alpha}^i_l \alpha^l_k = \delta^i_k \tag{55}$$

Expressed in the new coordinates the interval will again be a homogeneous quadratic form

$$ds^2 = g_{ik}\, dx^i dx^k = g'_{ik}\, dx'^i dx'^k, \tag{56}$$

where

$$g'_{ik} = g'_{ki} = g_{lm}\, \check{\alpha}^l_i\, \check{\alpha}^m_k \tag{57}$$

are new functions of the coordinates x'^i. The relations reciprocal to (57) are

$$g_{ik} = g'_{lm}\, \alpha^l_i\, \alpha^m_k, \tag{58}$$

which can easily be verified by substituting from (57) in the right-hand side of (58) and using (55). The transformations (54) must be such that the new functions g'_{ik} again satisfy inequalities of the form (52).

In general the system of reference R' defined by the system of coordinates (x'^i) will be different from the system of reference R corresponding to the coordinates (x^i), but if the transformations (54) are of the form

$$\left.\begin{aligned} x'^\iota &= x'^\iota(x^\kappa) \\ x'^4 &= x'^4(x^i) = f(x^i) \end{aligned}\right\}, \tag{59}$$

where the space coordinates x'^ι are functions of the spatial coordinates x^κ only, the systems of reference R' and R are identical. For in this case the transformation simply implies another notation for the points of reference in R together with an arbitrary continuous change in the rate and setting of the coordinate clocks. While each system of coordinates (x^i) corresponds to one and only one system of reference R, we can always in a given system of reference introduce an infinite number of different space-time coordinate systems which are connected by transformations of the form (59). The coefficients α^{ι}_{k} and $\breve{\alpha}^{\iota}_{k}$ corresponding to the special transformations (59) obviously satisfy the conditions

$$\alpha^{\iota}_{4} = \breve{\alpha}^{\iota}_{4} = 0. \tag{59'}$$

In general the gravitational fields in different systems of coordinates will be different, but for physical reasons it is convenient to consider the gravitational fields in all systems of coordinates connected by (59) as identical, since all these systems of coordinates correspond to the same system of reference. In different systems of reference, however, there will in general be different gravitational fields. In this respect the various systems of inertia are exceptional, since they all have a vanishing gravitational field.

89. Space and time measurements in an arbitrary system of reference. Experimental determination of the functions g_{ik}

Let us now consider an arbitrary system of reference R into which we have introduced a certain system of coordinates (x^{ι}). Consider in particular two points of reference A and B in this system with the space coordinates (x^{ι}) and $(x^{\iota}+dx^{\iota})$, respectively The spatial distance $d\sigma$ between A and B at the time $t = x^4/c$ can now be measured by means of a standard measuring-rod connecting the points A and B and at rest relative to A In the limit of very small dx^{ι}, B will also be practically at rest relative to the measuring-rod In order to express $d\sigma$ in terms of the functions g_{ik} we introduce the system of inertia I^0 relative to which the point of reference A (and approximately also B) is at rest at the time t If X^{ι} denote the pseudo-Cartesian space-time coordinates in I^0, the transformation from I^0 to R is given by the equations (42)–(48). However, since I^0 is the rest system of the point A at the time considered, we have, on account of (48),

$$\breve{\Lambda}^{\iota}_{4} = 0 \tag{60}$$

at the point A and the time t. The differences dX^{ι} between simultaneous values of the Cartesian coordinates of the points A and B at the time of the measurement are obtained from (44) by putting

$$i = \iota \quad \text{and} \quad dx^4 = dt = 0.$$

Hence
$$dX^{\iota} = \breve{\Lambda}^{\iota}_{\kappa}\, dx^{\kappa}. \tag{61}$$

On account of (60) the term $\breve{\Lambda}^{\iota}_{4}\, dx^4$ in (44) would vanish anyhow, even if $dx^4 \neq 0$, which means that the values of dX^{ι} would be approximately the same as for positions of A and B which are simultaneous relative to I^0.

According to the general assumption formulated on p. 223, the standard measuring-rod in the system R has the same length as the measuring-rod in I^0; therefore, $d\sigma^2$ is simply

$$d\sigma^2 = \sum_{\iota=1}^{3} (dX^\iota)^2.$$

Using the expression (61) for dX^ι we see that $d\sigma^2$ is a quadratic form in the differentials (dx^ι), i.e.

$$d\sigma^2 = \gamma_{\iota\kappa}\, dx^\iota dx^\kappa \tag{62}$$

with
$$\gamma_{\iota\kappa} = \bar{\mathrm{A}}_\iota^\lambda \bar{\mathrm{A}}_\kappa^\lambda.$$

Now we get from (47) and (60)

$$g_{\iota 4} = \bar{\mathrm{A}}_\iota^\iota \bar{\mathrm{A}}_4^\iota - \bar{\mathrm{A}}_\iota^4 \bar{\mathrm{A}}_4^4 = -\bar{\mathrm{A}}_\iota^4 \bar{\mathrm{A}}_4^4.$$

Hence
$$\bar{\mathrm{A}}_\iota^4 = -\frac{g_{\iota 4}}{\bar{\mathrm{A}}_4^4} = -\frac{g_{\iota 4}}{\sqrt{(-g_{44})}} = -\gamma_\iota,$$

where we have put
$$\gamma_\iota \equiv \frac{g_{\iota 4}}{\sqrt{(-g_{44})}}. \tag{63}$$

Furthermore, we obtain from (47) for $i = \iota$, $k = \kappa$

$$\bar{\mathrm{A}}_\iota^\lambda \bar{\mathrm{A}}_\kappa^\lambda = g_{\iota\kappa} + \bar{\mathrm{A}}_\iota^4 \bar{\mathrm{A}}_\kappa^4,$$

which leads to the following expressions for the spatial metric tensor

$$\gamma_{\iota\kappa} = g_{\iota\kappa} + \gamma_\iota \gamma_\kappa. \tag{64}$$

The metric tensor $\gamma_{\iota\kappa}$ which determines the spatial geometry in the reference system R will thus in general not simply be equal to the spatial part $g_{\iota\kappa}$ of the four-dimensional metric tensor g_{ik}. This is the case only if

$$g_{\iota 4} = 0 \quad \text{or} \quad \gamma_\iota = 0. \tag{65}$$

In a system of coordinates where the equations (65) are satisfied at every point in 4-space, the time axis is everywhere orthogonal to the spatial coordinate curves. Such a system will therefore be called *time-orthogonal.*

Now consider a standard clock C inserted at rest at a given reference point A of the system of reference R. The line element of the time track of this clock is then given by

$$ds^2 = g_{44}\,(dx^4)^2 \tag{66}$$

since $dx^\iota = 0$ for a clock at rest. While $t = x^4/c$ denotes the time shown by the coordinate clock at A, the increase $d\tau_0$ in the time of the clock C is given by

$$ds^2 = -c^2\, d\tau_0^2$$

(cf. § 37). This follows at once from the invariance of ds^2 if we introduce the system of inertia I^0 which is momentarily at rest relative to the clock C. Expressed in terms of the space-time coordinates (X^{0i}) of this system, ds^2 becomes

$$ds^2 = -c^2\,(dT^0)^2$$

since $dX^{0\iota} = 0$. Further, since the time T^0 itself is given by standard clocks at rest in I^0, we have $dT^0 = d\tau_0$, and (66) can be written

$$d\tau_0^2 = -g_{44}\,dt^2. \tag{66'}$$

Consequently the function $\sqrt{(-g_{44})}$ determines the ratio between the rates of the standard clock C and the coordinate clock at the place considered. The quantity g_{44} may thus be obtained experimentally by measuring the ratio of the rates of the two clocks in question. If we do this at each reference point and at all times, we get $g_{44}(x^\iota)$ as a function of the space-time coordinates (x^ι).

The metric tensor $\gamma_{\iota\kappa}$ which, by (62), determines the space geometry in S can now also be obtained by direct measurements, applying a method similar to that explained for the two-dimensional case in § 87. To determine the six independent components $\gamma_{\iota\kappa}$ we simply need to measure the lengths of six properly chosen line elements (dx^ι) at each point and at every time.

Now, if we could also find a procedure which would allow us to measure the quantity γ_ι, the metric tensor g_{ik} would be completely determined. To this purpose we consider a light signal which starts at the reference point A with space coordinates (x^ι) at the time t and arrives at a neighbouring point B: $(x^\iota+dx^\iota)$ at a time $t+dt$, say. The time track of this signal is characterized by the equation

$$ds^2 = g_{ik}\,dx^i dx^k = 0, \tag{67}$$

for in the system of inertia I the velocity of the signal is c, which means that the invariant

$$ds^2 = G_{ik}\,dX^i dX^k = dX^2+dY^2+dZ^2-c^2\,dT^2 = 0.$$

The equation (67) may also be written

$$g_{\iota\kappa}\,dx^\iota dx^\kappa+2g_{\iota 4}\,dx^\iota dx^4+g_{44}\,(dx^4)^2 = 0,$$

or, by means of (63), (64), and (62),

$$d\sigma^2-(\gamma_\iota\,dx^\iota)(\gamma_\kappa\,dx^\kappa)+2\sqrt{(-g_{44})}\gamma_\iota\,dx^\iota dx^4+g_{44}\,(dx^4)^2 = 0.$$

If we divide this equation by dt^2 we obtain

$$w^2 = \{\gamma_\iota w^\iota - c\sqrt{(-g_{44})}\}^2, \tag{68}$$

where the
$$w^\iota = \frac{dx^\iota}{dt} = wn^\iota \tag{69}$$

are the components of the velocity of the light signal in the direction n^ι and
$$w = \frac{d\sigma}{dt} = \sqrt{(\gamma_{\iota\kappa} w^\iota w^\kappa)} \tag{69'}$$

is the magnitude of this velocity. In the first place, we see from (68) that the light velocity w depends on the direction of propagation n^ι of the signal if $\gamma_\iota \neq 0$ in the system of coordinates considered.

In fact, using (69) in (68) we get
$$w(n^\iota) = \frac{c\sqrt{(-g_{44})}}{\gamma_\iota n^\iota + 1}, \tag{70}$$
where
$$\gamma_{\iota\kappa} n^\iota n^\kappa = 1, \quad \text{i.e. } (\gamma_\iota n^\iota)^2 < 1.$$

On the other hand, if we solve the equation (70) with respect to $(\gamma_\iota n^\iota)$, a measurement of the velocity of light in three different directions n^ι will allow us to determine the three quantities γ_ι. In this way it is possible in principle to determine all the quantities $\gamma_{\iota\kappa}$, γ_ι, g_{44}, i.e. the complete metric tensor g_{ik}, by experiments.

90. The spatial geometry in the rotating system of reference

Considering again the rotating system of coordinates introduced in § 84, we obtain from (1), (2), (3), and (10) the transformation equations (42), or rather the corresponding reciprocal equations, in the form
$$\left.\begin{array}{ll} X = r\cos(\vartheta+\omega t), & Y = r\sin(\vartheta+\omega t) \\ Z = z, & T = t \\ x^\iota = (r, \vartheta, z, ct) & \end{array}\right\}. \tag{71}$$

By differentiation and introduction of (dX, dY, dZ, dT) into the expression for the interval
$$ds^2 = dX^2+dY^2+dZ^2-c^2\,dT^2$$
we get
$$ds^2 = dr^2+r^2\,d\vartheta^2+dz^2+2\omega r^2\,d\vartheta dt-(c^2-r^2\omega^2)\,dt^2 = g_{ik}\,dx^i dx^k. \tag{72}$$

Thus we have
$$\left.\begin{array}{llll} g_{11} = 1, & g_{22} = r^2, & g_{33} = 1, & g_{44} = -\left(1-\dfrac{r^2\omega^2}{c^2}\right) \\ g_{24} = g_{42} = \dfrac{\omega r^2}{c} & & & \end{array}\right\}, \tag{73}$$

all other components of $g_{\iota k}$ being zero. Hence, from (63) and (64),

$$\left.\begin{aligned} \gamma_\iota &= \left(0, \frac{\omega r^2}{\sqrt{(c^2-r^2\omega^2)}}, 0\right) = \frac{\omega r^2}{\sqrt{(c^2-r^2\omega^2)}}\,\delta_{\iota 2} \\ \gamma_{11} &= 1, \qquad \gamma_{22} = \frac{r^2}{1-r^2\omega^2/c^2}, \qquad \gamma_{33} = 1 \\ \gamma_{\iota\kappa} &= 0 \text{ for } \iota \neq \kappa \end{aligned}\right\}. \tag{74}$$

Thus if we calculate the spatial line-element $d\sigma^2$ by means of (62) and (74) we come back to the expression (7). Further, (66′) together with the expression (73) for g_{44} leads back to the equation (9).

The space geometry defined by the line-element $d\sigma^2 = \gamma_{\iota\kappa}\,dx^\iota dx^\kappa$ is non-Euclidean. In the plane $z = Z = 0$, i.e. on the rotating disk, we have only the two coordinates $(x^1, x^2) = (r, \vartheta)$ and the geodesics are determined by the equations (32)

$$\frac{d}{d\sigma}\left(\gamma_{\iota\kappa}\frac{dx^\kappa}{d\sigma}\right) = \frac{1}{2}\frac{\partial\gamma_{\kappa\lambda}}{\partial x^\iota}\frac{dx^\kappa}{d\sigma}\frac{dx^\lambda}{d\sigma} \quad (\iota = 1, 2), \tag{75a}$$

$$\gamma_{\iota\kappa}\frac{dx^\iota}{d\sigma}\frac{dx^\kappa}{d\sigma} = 1 \tag{75b}$$

They define the curves of shortest distance as measured by measuring-rods at rest on the rotating disk. In our case we get from the second equation (75a) and (74)

$$\frac{d}{d\sigma}\left(\frac{r^2}{1-r^2\omega^2/c^2}\frac{d\vartheta}{d\sigma}\right) = 0,$$

and by integration

$$\frac{r^2}{1-r^2\omega^2/c^2}\dot{\vartheta} = \alpha, \tag{76a}$$

where α is a constant, and $\dot{\vartheta} = d\vartheta/d\sigma$. Hence,

$$\dot{\vartheta} = \alpha\left(\frac{1-r^2\omega^2/c^2}{r^2}\right). \tag{76a'}$$

Further, from (75b)

$$\dot{r}^2 + \frac{r^2}{1-r^2\omega^2/c^2}\dot{\vartheta}^2 = 1, \qquad \dot{r} = \pm\sqrt{\left(1+\frac{\alpha^2\omega^2}{c^2}-\frac{\alpha^2}{r^2}\right)}. \tag{76b}$$

Hence,

$$\frac{dr}{d\vartheta} = \frac{\dot{r}}{\dot{\vartheta}} = \pm\frac{r^2\sqrt{\left(1+\frac{\alpha^2\omega^2}{c^2}-\frac{\alpha^2}{r^2}\right)}}{\alpha\left(1-\frac{r^2\omega^2}{c^2}\right)}, \tag{77}$$

which by integration gives r as a function of ϑ for the geodesics.

If the constant of integration α is zero, we get from (76 a') $\dot{\vartheta} = 0$, i.e. the radius vectors with $\vartheta =$ constant are geodesics. In Fig. 17 we have given a picture in a Euclidean plane of some of the geodesics on the rotating disk. In this picture the points on the disk with the coordinates (r, ϑ) are depicted as points with the polar coordinates (r, ϑ). The curves

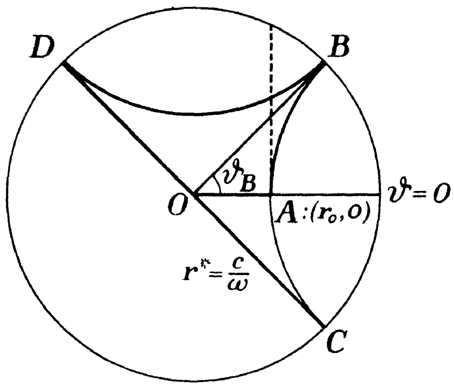

FIG 17

in this picture are therefore simply the curves in the fixed XY-plane with which the points on the geodesics of the rotating disk coincide at the time $t = T = 0$. The radius r^* of the disk is defined by (4), i.e.

$$r^* = c/\omega. \tag{78}$$

For the geodesics OB and OA we have $\alpha = 0$. The geodesic AB which starts at the point $(r_0, 0)$ at right angles to OA, i.e. with $\dot{r} = 0$, corresponds to an $\alpha \neq 0$ which follows from (76 a'), (76 b):

$$\alpha = \frac{r_0^2}{(1-r_0^2\,\omega^2/c^2)}\,\dot{\vartheta}_0 = \frac{r_0}{\sqrt{(1-r_0^2\,\omega^2/c^2)}}. \tag{79}$$

With increasing values of ϑ the radius vector r increases also with a rate determined by (77), and it is seen at once that $dr/d\vartheta$ is larger than in the limit $c \to \infty$, where we should have

$$\frac{dr}{d\vartheta} = \frac{r(r^2-\alpha^2)^{\frac{1}{2}}}{\alpha}. \tag{80}$$

The broken straight line in Fig. 17 is the curve determined by (80) which would be the geodesic if the geometry on the rotating disk were Euclidean. When r approaches the value $r^* = c/\omega$ we get $\dot{\vartheta} \to 0$ from (76 a'), which means that the geodesic AB has OB as tangent at the point B.

Consider two geodesics $x_1^i = x_1^i(\sigma_1)$, $x_2^i = x_2^i(\sigma_2)$ which go through the same point. The angle θ between the two curves at the point of intersection is determined by (19). Thus, on account of (74) and (76),

$$\cos\theta = \frac{\gamma_{\iota\kappa}\, dx_1^\iota\, dx_2^\kappa}{d\sigma_1\, d\sigma_2} = \dot{r}_1\dot{r}_2 + \frac{r^2\dot{\vartheta}_1\dot{\vartheta}_2}{1-r^2\omega^2/c^2}$$

$$= \pm\sqrt{\left(1+\frac{\alpha_1^2\omega^2}{c^2}-\frac{\alpha_1^2}{r^2}\right)}\sqrt{\left(1+\frac{\alpha_2^2\omega^2}{c^2}-\frac{\alpha_2^2}{r^2}\right)}+\frac{\alpha_1\alpha_2(1-r^2\omega^2/c^2)}{r^2}, \tag{81}$$

where α_1 and α_2 are the values of the constants α for the two geodesics in question. This expression can be used everywhere except at the point O.

Now consider the triangle OAB. At O the angle between the two sides of the triangle is equal to $\theta_O = \vartheta_B < \frac{1}{2}\pi$, for in the centre of the disk the geometry is Euclidean. The angle between the geodesics AO and AB at A is found from (81) by putting

$$\alpha_1 = 0, \qquad \alpha_2 = \alpha = \frac{r_0}{\sqrt{(1-r_0^2\,\omega^2/c^2)}};$$

thus,

$$\cos\theta_A = -\sqrt{\left\{1+\alpha^2\left(\frac{\omega^2}{c^2}-\frac{1}{r_0^2}\right)\right\}} = 0,$$

i.e.

$$\theta_A = \tfrac{1}{2}\pi.$$

Finally, we get the angle θ_B between the geodesics BA and BO from (81) by putting $\alpha_1 = 0$, $\alpha_2 = -\alpha$, $r = r^* = c/\omega$. Thus we get

$$\cos\theta_B = \sqrt{\left\{1+\alpha^2\left(\frac{\omega^2}{c^2}-\frac{\omega^2}{c^2}\right)\right\}} = 1, \qquad \theta_B = 0.$$

Hence the sum of the angles in the triangle OAB is

$$\theta_{OAB} = \theta_O+\theta_A+\theta_B = \vartheta_B+\tfrac{1}{2}\pi < \pi,$$

i.e. the sum of the angles in a triangle on the rotating disk is smaller than π. Only if the triangle is close to the centre is the sum of the angles approximately equal to the Euclidean value π. There are even triangles on the rotating disk for which the sum of the angles is zero. If, for instance, $\vartheta_B = \frac{1}{4}\pi$, we see that the triangle CBD formed by the geodesics CB, BD, and DC has a vanishing sum, i.e.

$$\theta_{CBD} = 0.$$

Thus the sum of the angles in a triangle on the rotating disk can have all values between 0 and π. The spatial geometry as determined by observers on the rotating disk is the same as on a surface of negative curvature in a three-dimensional Euclidean space.

Sometimes it is convenient to use a different set of space coordinates (x, y, z) in the rotating system connected with the cylindrical coordinates (r, ϑ, z) by the equations

$$x = r\cos\vartheta, \qquad y = r\sin\vartheta, \qquad r^2 = x^2+y^2. \tag{82}$$

In these coordinates the line element (72) takes the form

$$ds^2 = dx^2+dy^2+dz^2+2\omega(-y\,dx+x\,dy)\,dt-\left(1-\frac{r^2\omega^2}{c^2}\right)c^2\,dt^2. \tag{83}$$

91. The time tracks of free particles and light rays

Consider a material particle which is moving freely under the influence solely of the gravitational fields in an accelerated system with the coordinates (x^ι). Since these fields are supposed to be non-permanent, they can be transformed away simply by introducing the pseudo-Cartesian coordinates (X^ι) of the system of inertia I from which we started in § 88. In this system the motion of a free particle is uniform, i.e. its time track is a straight line defined by the equation

$$\frac{d^2X^\iota}{d\lambda^2} = 0, \tag{84}$$

where λ is an arbitrary parameter. In other words, the time track of a freely falling particle is a geodesic in 4-space. The geodesics are defined by the variational principle (21), (29), (33), where the indices ι, k now are running from 1 to 4. If we introduce the proper time $\tau = s/ic$ as parameter, the variational principle (33) states that the variation

$$\delta\int\limits_{\tau_1}^{\tau_2} d\tau = \delta\int\limits_{\tau_1}^{\tau_2}\frac{1}{c}\sqrt{\left(-g_{\iota k}\frac{dx^\iota}{d\tau}\frac{dx^k}{d\tau}\right)}\,d\tau$$
$$= \frac{1}{c^2}\int\left[\frac{d}{d\tau}\left(g_{\iota k}\frac{dx^k}{d\tau}\right)-\frac{1}{2}\frac{\partial g_{kl}}{\partial x^\iota}\frac{dx^k}{d\tau}\frac{dx^l}{d\tau}\right]\delta x^\iota d\tau \tag{85}$$

must be zero for all variations δx^ι which vanish for $\tau = \tau_1$ and $\tau = \tau_2$. This leads to the Euler equations (32) or

$$\frac{d}{d\tau}\left(g_{\iota k}\frac{dx^k}{d\tau}\right) = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^\iota}\frac{dx^k}{d\tau}\frac{dx^l}{d\tau}, \qquad g_{\iota k}\frac{dx^\iota}{d\tau}\frac{dx^k}{d\tau} = -c^2. \tag{86}$$

In a pseudo-Cartesian system of coordinates the metric tensor is $G_{\iota k}$ as defined by (41). The equations (86) then reduce to (84), and (85) reduces to the equation (V. 102) used in § 59.

Now consider a light ray in empty space. In the system of inertia I its time track is again given by (84), but with the extra condition that $ds^2 = G_{\iota k}\,dX^\iota dX^k = 0$. The time track of a light ray is thus a geodesic of zero length, and we can therefore not use the length as a parameter. Instead of (86) we thus have in an arbitrary system of coordinates (x^i) the equations (30) and (31) with the constant in (31) equal to zero, i.e.

$$\frac{d}{d\lambda}\left(g_{ik}\frac{dx^k}{d\lambda}\right) = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda}, \qquad g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda} = 0, \tag{87}$$

where

$$x^i = x^i(\lambda)$$

may be any parametric representation of the time track.

92. The dynamical gravitational potentials

In §§ 89 and 90 we have seen that the gravitational field in an arbitrary accelerated system of coordinates (x^i) influences the space and time measurements as made by standard measuring-rods and standard clocks. The space geometry, for example, is determined by the spatial tensor $\gamma_{\iota\kappa}$ defined by (64). We shall now determine the quantities which describe the dynamical action of the gravitational fields. For this purpose we shall use a test body of arbitrary mass, which is placed at rest at the point of our system of reference at which we wish to measure the gravitational field. The acceleration imparted to this particle by the gravitational field then determines the strength of the field.

From (86) we get, for a particle which is momentarily at rest, taking $i = \iota = 1, 2, 3$,

$$g_{\iota\kappa}\frac{d^2x^\kappa}{d\tau^2} + \frac{d}{d\tau}\left(g_{\iota 4}\frac{dx^4}{d\tau}\right) = \frac{1}{2}\frac{\partial g_{44}}{\partial x^\iota}\left(\frac{dx^4}{d\tau}\right)^2. \tag{88}$$

Further, we get by means of (64), (63), and (62) for a particle in arbitrary motion

$$\begin{aligned}-c^2\,d\tau^2 = ds^2 &= g_{\iota\kappa}\,dx^\iota dx^\kappa + 2g_{\iota 4}\,dx^\iota dx^4 + g_{44}\,(dx^4)^2 \\ &= \gamma_{\iota\kappa}\,dx^\iota dx^\kappa + \frac{(g_{\iota 4}\,dx^\iota + g_{44}\,dx^4)^2}{g_{44}} \\ &= d\sigma^2 + g_{44}\,(dx^4)^2\left(1 - \frac{\gamma_\iota}{\sqrt{(-g_{44})}}\frac{u^\iota}{c}\right)^2,\end{aligned} \tag{89}$$

where τ is the proper time, and the

$$u^\iota = \frac{dx^\iota}{dt} = c\frac{dx^\iota}{dx^4}$$

are the components of the velocity of the particle.

Dividing (89) by $(dx^4)^2$ and solving this equation with respect to $dx^4/d\tau$ gives

$$\frac{dx^4}{d\tau} = c\left\{(-g_{44})\left(1 - \frac{\gamma_\iota u^\iota}{c\sqrt{(-g_{44})}}\right)^2 - u^2/c^2\right\}^{-\frac{1}{2}}, \tag{90}$$

where
$$u = \frac{d\sigma}{dt}$$
is the magnitude of the particle velocity.

Hence
$$g_{\iota 4}\frac{dx^4}{d\tau} = c\gamma_\iota\left\{\left(1-\frac{\gamma_\kappa u^\kappa}{c\sqrt{(-g_{44})}}\right)^2 + u^2/c^2 g_{44}\right\}^{-\frac{1}{2}},$$
and for a particle which is momentarily at rest we get after some calculation
$$\left.\begin{aligned}\frac{d}{d\tau}\left(g_{\iota 4}\frac{dx^4}{d\tau}\right) &= \frac{1}{\sqrt{(-g_{44})}}\left[c^2\frac{\partial\gamma_\iota}{\partial x^4}+\frac{\gamma_\iota\gamma_\kappa}{\sqrt{(-g_{44})}}\frac{d^2x^\kappa}{dt^2}\right]\\ \text{and}\qquad \frac{d^2x^\kappa}{d\tau^2} &= \frac{1}{-g_{44}}\frac{d^2x^\kappa}{dt^2}\end{aligned}\right\}. \tag{91}$$
Using (91) and (90) in (88) we thus obtain
$$\gamma_{\iota\kappa}\frac{d^2x^\kappa}{dt^2} = -\frac{\partial}{\partial x^\iota}\left(-\frac{c^2 g_{44}}{2}\right) - c\sqrt{(-g_{44})}\frac{\partial\gamma_\iota}{\partial t}. \tag{92}$$
The left-hand side represents the gravitational acceleration imparted to the test particle,
$$a_\iota = \gamma_{\iota\kappa}\frac{d^2x^\kappa}{dt^2} \tag{93}$$
being the covariant components of the acceleration expressed in the curvilinear coordinates of our system of reference (see § 100). Thus the dynamical action of the gravitational field is described by the functions g_{44} and γ_ι.

If we put
$$g_{44} = -\left(1+\frac{2\chi}{c^2}\right), \tag{94}$$
we get from (92)
$$a_\iota = -\frac{\partial\chi}{\partial x^\iota} - c\sqrt{\left(1+\frac{2\chi}{c^2}\right)}\frac{\partial\gamma_\iota}{\partial t}. \tag{95}$$

On account of the analogy of this expression with the expression for the electric force on a charged particle at rest in terms of the electromagnetic potentials, the quantities χ and γ_ι will be called the gravitational scalar and vector potential, respectively. The scalar potential χ has been normalized so as to give g_{44} the value -1 of the special theory of relativity for a vanishing potential.

If γ_ι is time-independent, the gravitational acceleration is simply equal to the gradient of the scalar potential:
$$a_\iota = -\frac{\partial\chi}{\partial x^\iota}. \tag{96}$$

This is, for instance, the case in the rotating system of coordinates considered in § 90, and from (73) and (94) we get in this case

$$\chi = -\tfrac{1}{2}r^2\omega^2. \tag{97}$$

Thus the gravitational acceleration of a particle of zero velocity lies in the direction of increasing r and is equal to $r\omega^2$. This is in accordance with the usual expression for the centrifugal force.

93. The rate of a moving standard clock in a gravitational field

From (90) we get for the proper time of a particle moving in the gravitational field described by the potentials (γ_ι, χ)

$$d\tau = dt\left\{\left[\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\iota u^\iota}{c}\right]^2 - \frac{u^2}{c^2}\right\}^{\frac{1}{2}}. \tag{98}$$

Since τ is the time measured by a standard clock following the particle in its motion, (98) gives the rate of the moving standard clock compared with the rate dt of the coordinate clocks of the system considered. If the system of coordinates is time-orthogonal, we have $\gamma_\iota = 0$ and

$$d\tau = dt\left(1+\frac{2\chi}{c^2}-\frac{u^2}{c^2}\right)^{\frac{1}{2}}. \tag{99}$$

The formulae (98) and (99) are the generalizations of the formula (II. 38) and express the retardation (or advancement) of moving clocks in the case where gravitational fields are present.

For a clock at rest in our system of reference we have

$$d\tau_0 = dt\sqrt{(1+2\chi/c^2)} \tag{100}$$

in accordance with (66′) and (94). The rate of a standard clock thus depends on the scalar gravitational potential at the place where the clock is situated; it is lower at places of small gravitational potential.

On the rotating disk we have, according to (100) and (97),

$$d\tau_0 = dt\sqrt{(1+2\chi/c^2)} = dt\sqrt{(1-r^2\omega^2/c^2)}, \tag{101}$$

i.e. a standard clock far from the centre has a slower rate than a standard clock placed at the centre which simply shows the time $t = T$. This retardation of a standard clock at a place with $r > 0$ will be differently interpreted by observers in the fixed system I and in the rotating system S. An observer in I will explain the retardation by the motion of the particle. In this system we have no gravitational field, i.e. $\chi = 0$, but the velocity of the clock is $u = r\omega$. An application of (99) in I thus leads to the formula

$$d\tau_0 = dt\sqrt{(1-u^2/c^2)} = dt\sqrt{(1-r^2\omega^2/c^2)}. \tag{102}$$

On the other hand, in S the velocity $u = 0$, but we have a gravitational field $\chi = -\frac{1}{2}r^2\omega^2$, and (99) again leads to the same expression (101) or (102). An observer in S will thus explain the retardation by the action of the gravitational field present in the rotating system.

94. Transformation of coordinates inside a fixed system of reference

Let (x^i) be a set of space-time coordinates corresponding to a certain system of reference R. By the transformation (59) we may then introduce new space-time coordinates inside the same system of reference R. The transformation (59) simply introduces a new numbering of the reference points in R together with an arbitrary change in rate and setting of the coordinate clocks This, of course, cannot give rise to any change in the spatial geometry in R determined by measurements with standard measuring-rods, i.e. $d\sigma$ as defined by (62), (64), and (63) must be invariant by the transformation (59). A formal proof of this statement is given in Appendix 4.

Since the system of reference is unchanged by the transformation (59), the gravitational field must also be regarded as unchanged. The gravitational potentials (χ, γ_ι) defined by (94) and (63) will, however, be transformed in accordance with (57). In this respect a transformation of that kind is thus analogous to the gauge transformation (V. 23) of the electromagnetic potentials by which these potentials are changed without any influence on the electromagnetic field derived from the potentials. In many cases it is possible by such a 'gauge transformation' of the gravitational potentials to give the potentials a particularly simple form. In the first place, it is always possible by a transformation of the type

$$x'^\iota = x^\iota; \qquad x'^4 = x'^4(x^i) = f(x^i) \tag{103}$$

to ensure that the scalar potential in the system (x'^i) vanishes. We need only choose the new time variable $t' = x'^4/c$ so that the new coordinate clocks have the same rate as standard clocks at rest. On account of (66′) and (100), this is obtained by putting

$$t' = \tau_0 = \int\limits_0^t \sqrt{(-g_{44})}\, dt + \psi(x^\iota) = \int\limits_0^t \sqrt{\left(1+\frac{2\chi}{c^2}\right)}\, dt + \psi(x^\iota), \tag{104}$$

where the integration is performed for constant values of the space coordinates (x^ι), and $\psi(x^\iota)$ may be any function of the space coordinates. With this choice of the new time variable we get, since (100) must hold also in the new system,

$$dt' = d\tau_0 = dt'\sqrt{\left(1+\frac{2\chi'}{c^2}\right)} = dt'\sqrt{(-g'_{44})},$$

i.e.
$$\chi' = 0, \qquad g'_{44} = -1. \tag{105}$$

A more useful simplification would, however, be obtained if the vector potential could be 'transformed away' by a transformation of the type considered, for this would mean that the new system of coordinates is time-orthogonal and all formulae are considerably reduced Let us therefore try to find a transformation (103) such that

$$\gamma'_\iota = 0. \tag{106}$$

Since the interval can be written in the form (89) in every system of coordinates, and since $d\sigma^2$ is invariant under the transformations (103), we must have

$$g'_{44}(dx'^4)^2 = \left[dx^4 - \frac{\gamma_\iota}{\sqrt{(-g_{44})}}\,dx^\iota\right]^2 g_{44},$$

i e
$$dx'^4 = \pm\left(\frac{g_{44}}{g'_{44}}\right)^{\frac{1}{2}}\left(dx^4 - \frac{\gamma_\iota\,dx^\iota}{\sqrt{(-g_{44})}}\right) \tag{107}$$

must be the total differential of the function $f(x^\iota)$ in (103). Thus we see that the space-time derivatives of the function f must have the following ratios.

$$\frac{\partial f}{\partial x^1}:\frac{\partial f}{\partial x^2}:\frac{\partial f}{\partial x^3}:\frac{\partial f}{\partial x^4} = -\frac{\gamma_1}{\sqrt{(-g_{44})}}:-\frac{\gamma_2}{\sqrt{(-g_{44})}}:-\frac{\gamma_3}{\sqrt{(-g_{44})}}:1. \tag{108}$$

This is equivalent to the validity of the three equations

$$\frac{\partial f}{\partial x^\iota}+\frac{\gamma_\iota}{\sqrt{(1+2\chi/c^2)}}\frac{\partial f}{\partial x^4} = 0 \quad \text{for} \quad \iota = 1,2,3. \tag{109}$$

These conditions could also be obtained from (106) by means of the transformation properties of the gravitational potentials (see Appendix 4). Now the simultaneous differential equations (109) have a solution only when the following compatibility conditions are satisfied. If we multiply (109) by the operator

$$\frac{\partial}{\partial x^\kappa}+\frac{\gamma_\kappa}{\sqrt{(1+2\chi/c^2)}}\frac{\partial}{\partial x^4}$$

and subtract the equation obtained in this way from the equation with ι

and κ interchanged, we get after a simple calculation the condition that the quantity

$$\omega_{\iota\kappa} \equiv \left(\frac{\partial}{\partial x^\iota} + \frac{\gamma_\iota}{\sqrt{(1+2\chi/c^2)}} \frac{\partial}{\partial x^4}\right) \frac{\gamma_\kappa}{\sqrt{(1+2\chi/c^2)}} - \\ - \left(\frac{\partial}{\partial x^\kappa} + \frac{\gamma_\kappa}{\sqrt{(1+2\chi/c^2)}} \frac{\partial}{\partial x^4}\right) \frac{\gamma_\iota}{\sqrt{(1+2\chi/c^2)}} \tag{110}$$

must be zero, i.e.

$$\omega_{\iota\kappa} = 0 \tag{111}$$

for all values of ι and κ. (111) together with (110) represent the general condition which the dynamical gravitational potentials must satisfy in order that the vector potential may be made to vanish by a transformation of the type (103).†

In the case of the rotating system of coordinates considered in § 90, we see at once from (110), (73), and (74) that the only non-vanishing components of $\omega_{\iota\kappa}$ are

$$\omega_{12} = -\omega_{21} = \frac{2r\omega}{c(1-r^2\omega^2/c^2)^2}, \tag{112}$$

which are different from zero for $r > 0$. The condition (111) is therefore not fulfilled in the rotating system and it is *not* possible by a simple change of rate of the coordinate clocks to introduce a time-orthogonal system of coordinates in this system of reference.

The gravitational field in a given system of reference R is called *stationary* if it is possible, by a suitable choice of the space-time coordinates in R, to ensure that all components of g_{ik}, i.e. $\gamma_{\iota\kappa}$, χ, and γ_ι, are independent of the time variable. If simultaneously we can obtain $\gamma_\iota = 0$, the gravitational field is called *static*. The gravitational field in the uniformly rotating system of reference is thus stationary.

95. Further simple examples of accelerated systems of reference

Let $X^\iota = (X, Y, Z, cT)$ again be pseudo-Cartesian space-time coordinates in a system of inertia I. The Galilean transformation (I. 2)

$$x = X - vT, \qquad y = Y, \qquad z = Z, \qquad t = T \tag{113}$$

then defines a new system of reference which obviously is the system of inertia I' moving in the direction of the X-axis with the velocity v relative to I; for each reference point $(x, y, z) =$ constant is moving with the

† This condition was found by Weyssenhoff, see J. v. Weyssenhoff, *Bull. Acad. Polonaise*, Ser. A, p. 252 (1937); see also *ZS. f. Phys.* **95**, 391 (1935); ibid. **107**, 64 (1937).

same velocity v. If we put $(x^\iota) = (x, y, z, ct)$ the interval takes the form

$$\begin{aligned} ds^2 &= g_{\iota k}\, dx^\iota dx^k = G_{ik}\, dX^\iota dX^k \\ &= dx^2+dy^2+dz^2+2v\, dxdt-c^2\, dt^2(1-v^2/c^2), \end{aligned} \tag{114}$$

i.e. $g_{11} = g_{22} = g_{33} = 1,\quad g_{14} = g_{41} = v/c,\quad g_{44} = -(1-v^2/c^2),$

all other components vanishing. The system of coordinates (x^ι) is therefore not time-orthogonal and from (63), (64), and (94) we get

$$\left.\begin{aligned} &\gamma_\iota = \left(\frac{v}{\sqrt{(c^2-v^2)}}, 0, 0\right), \qquad \chi = -\frac{v^2}{2} \\ &\gamma_{11} = \frac{1}{1-v^2/c^2}, \qquad \gamma_{22} = \gamma_{33} = 1, \qquad \gamma_{\iota\kappa} = 0 \text{ for } \iota \neq \kappa \end{aligned}\right\}. \tag{115}$$

Since the gravitational potentials are constant the gravitational field determined by (95) is, of course, zero in I', and the potentials may be transformed away by a transformation (59) with f of the form (104). Introducing new coordinates $X'^\iota = (X', Y', Z', cT')$ by

$$\left.\begin{aligned} &X' = \frac{x}{\sqrt{(1-v^2/c^2)}}, \qquad Y' = y, \qquad Z' = z \\ &T' = \sqrt{(1-v^2/c^2)}\, t - \frac{vx}{c^2\sqrt{(1-v^2/c^2)}} \end{aligned}\right\}, \tag{116}$$

we get

$$\left.\begin{aligned} &ds^2 = dX'^2+dY'^2+dZ'^2-c^2\, dT'^2 \\ &g'_{\iota k} = G_{\iota k} \\ &\gamma'_{\iota\kappa} = \delta_{\iota\kappa}, \quad \gamma'_\iota = 0, \quad \chi' = 0 \end{aligned}\right\}. \tag{117}$$

The coordinates (X'^ι) are pseudo-Cartesian coordinates in I', they are connected with the (X^ι) by the special Lorentz transformation.

As a further example we consider the accelerated system

$$(x^\iota) = (x, y, z, ct)$$

defined by the transformation

$$X = x+\tfrac{1}{2}gt^2, \qquad Y = y, \qquad Z = z, \qquad T = t. \tag{118}$$

Each reference point in the system (x^ι) has a constant acceleration g in the direction of the X-axis relative to I. A simple calculation gives

$$ds^2 = dx^2+dy^2+dz^2+2gt\, dxdt-c^2\, dt^2\left(1-\frac{g^2t^2}{c^2}\right), \tag{119}$$

i.e. $g_{11} = g_{22} = g_{33} = 1,\quad g_{14} = g_{41} = gt/c,\quad g_{44} = -\left(1-\frac{g^2t^2}{c^2}\right),$

all other components $g_{\iota k}$ being zero.

Hence,

$$\left.\begin{aligned} \gamma_\iota &= \left(\frac{gt}{c\sqrt{(1-g^2t^2/c^2)}}, 0, 0\right), \qquad \chi = -\frac{g^2t^2}{2} \\ \gamma_{11} &= \frac{1}{1-g^2t^2/c^2}, \qquad \gamma_{22} = \gamma_{33} = 1, \qquad \gamma_{\iota\kappa} = 0 \text{ for } \iota \neq \kappa \end{aligned}\right\}. \tag{120}$$

This system of coordinates corresponds to a physically realizable system of reference only for $t = T < c/g$, for only in this case will the velocity gt of the reference points relative to I be smaller than c and $g_{44} < 0$.

The space geometry in this system is defined by the spatial line element

$$d\sigma^2 = \gamma_{\iota\kappa}\, dx^\iota dx^\kappa = \frac{dx^2}{1-g^2t^2/c^2} + dy^2 + dz^2. \tag{121}$$

It is non-Euclidean on account of the Lorentz contraction of the measuring-rods in the moving system. Since the potentials do not depend on the space coordinates the quantities $\omega_{\iota\kappa}$ in (110) are zero, which means that the vector potentials can be made equal to zero by a transformation (103). Integrating the equations (109) we find that the transformation

$$x' = x, \qquad y' = y, \qquad z' = z, \qquad t' = te^{\,g(x+\frac{1}{2}gt^2)/c^2} \tag{122}$$

leads to the desired result, i.e. if ds^2 is written in the form (89) we get in the new coordinates

$$ds^2 = \frac{dx'^2}{1-g^2t^2/c^2} + dy'^2 + dz'^2 + g'_{44}\, c^2\, dt'^2, \tag{122'}$$

with

$$\left.\begin{aligned} g'_{44} &= -\frac{e^{2g(x+\frac{1}{2}gt^2)/c^2}}{1-g^2t^2/c^2} = -\frac{e^{2gX/c^2}}{1-g^2T'^2/c^2} \\ \gamma'_\iota &= 0, \qquad \gamma'_{\iota\kappa} = \gamma_{\iota\kappa}, \qquad \chi' = -\frac{c^2(g'_{44}+1)}{2} \end{aligned}\right\}. \tag{123}$$

It is easily verified by direct calculation that (122′) is identical with (119) when use is made of the transformation equations (122).

The gravitational field in the accelerated system considered is non-static. According to (121) even the space geometry is time-dependent, i.e. the distance between two neighbouring reference points depends on the variable t. This is also evident, since the measuring-rods in the accelerated system on account of the increasing velocity gt relative to I will undergo an increasing Lorentz contraction. Although our system of reference is rigid from the point of view of an observer in I, an observer in the accelerated system itself will find that the system of reference points is dilated in the direction of the x-axis.

For small values of t and x, i.e. if we retain only terms of the first order in gt/c and gx/c^2, we get $t' = t(1-gx/c^2)$, $g'_{44} = -(1+2gx'/c^2)$, $\chi' = gx'$, and from (96)

$$a'_\iota = -\frac{\partial \chi'}{\partial x'^\iota} = (-g, 0, 0), \tag{124}$$

i.e. the gravitational field is constant in this region.

96. Rigid systems of reference with an arbitrary motion of the origin

A system of reference is called *rigid* if the distance between two reference points, as measured by standard measuring-rods at rest in the system, is constant in time. Thus the uniformly rotating system discussed in § 90 is rigid, while the system considered at the end of § 95 is not.

Consider now a particle in arbitrary motion relative to the system I with the coordinates $X_\iota = (X, Y, Z, icT)$. Its time track may be described by the equations

$$X_\iota = f_\iota(\tau), \tag{125}$$

τ being the proper time of the particle. We shall now try to introduce a system of coordinates $(x^\iota) = (x, y, z, ct)$ which is the relativistic analogue of a classical rigid frame of Cartesian axes following the particle in its motion, so that the particle is constantly situated at the origin of this frame of reference, and the space axes have constant directions. In § 46 we have determined the successive systems of inertia $S'(\tau)$ which are momentary rest systems of the particle and which are successively obtained by infinitesimal Lorentz transformations without rotation of the spatial axes. The transformation from the fixed system (X_ι) to the coordinates x'_ι in $S'(\tau)$ is then given by (IV. 140), where the coefficients $\alpha_{ik}(\tau)$ are determined by the differential equations (IV. 136, 137).

Now defining the system (x^ι) by putting

$$x^\iota = x'_\iota, \qquad x'_4 = 0, \qquad t = \tau \tag{126}$$

in (IV. 140), we find that simultaneous positions of the reference points in the system (x^ι) at the time t coincide with the simultaneous positions of the reference points in $S'(\tau)$ at the time $x'_4 = 0$, and the coinciding reference points have the same values for the spatial coordinates in these two systems of coordinates. The transformation connecting the variables (X_ι) and (x^ι) are thus

$$X_\iota = f_\iota(t) + x^\kappa \alpha_{\kappa\iota}(t), \tag{127}$$

where the coefficients α_{ik} are the solutions of (IV. 136). These equations are completely determined by the given motion (125) of the particle which lies permanently at the origin $x^\iota = 0$ of the system (x^ι). Differentiation of (127) gives, by means of (IV. 136),

$$\left.\begin{aligned} dX_i &= [\dot{f}_\iota(t)+x^\kappa\dot{\alpha}_{\kappa\iota}(t)]\,dt+dx^\kappa\alpha_{\kappa\iota}(t) \\ &= (U_i+x^\kappa\alpha_{\kappa l}\,\eta_{li})\,dt+dx^\kappa\alpha_{\kappa\iota} \end{aligned}\right\}. \tag{128}$$

If we use (128) in the expression for the interval we get, on account of (IV. 138),

$$\left.\begin{aligned} ds^2 &= dX_\iota\, dX_\iota = dx^2+dy^2+dz^2-c^2\,dt^2\left[1+\frac{g_\kappa x^\kappa}{c^2}\right]^2 \\ x^\iota &= (x^\iota, ct) = (x,y,z,ct) \end{aligned}\right\}, \tag{129}$$

where
$$g_\kappa = g_\kappa(t) \doteq \alpha_{\kappa l}\,\dot{U}_l = \alpha_{\kappa l}\,\ddot{f}_l \tag{130}$$

are functions of t only, which are completely determined by the motion of the origin of our system of coordinates (x^ι) relative to the system (X_ι). The quantities g_κ are equal to the components of the acceleration of the particle in its momentary rest system $S'(t)$ (cf. (IV. 42, 42′)).

The system of coordinates (x^ι) defined by (127) is time-orthogonal. The corresponding system of reference is rigid, for the distance between two reference points (x,y,z) and $(x+dx, y+dy, z+dz)$ is given by

$$d\sigma^2 = dx^2+dy^2+dz^2. \tag{131}$$

Thus the space geometry is even Euclidean and (x,y,z) are Cartesian space coordinates. The vector potential is zero and for the scalar potential we get

$$\chi = -\frac{c^2}{2}(g_{44}+1) = \frac{c^2}{2}\left[\left(1+\frac{g_\iota x^\iota}{c^2}\right)^2-1\right] = (\mathbf{g}\,.\,\mathbf{x})\left(1+\frac{(\mathbf{g}\,.\,\mathbf{x})}{2c^2}\right). \tag{132}$$

Its dependence on the space variables is thus always of the same type, only the coefficients $\mathbf{g} = \mathbf{g}(t)$ will be different for different motions of the origin. The gravitational field is determined by (96) and (132):

$$\mathbf{a} = -\mathrm{grad}\,\chi = -\mathbf{g}(1+(\mathbf{g}\,.\,\mathbf{x})/c^2). \tag{133}$$

Hence, in a region around the origin, where

$$(\mathbf{g}\,.\,\mathbf{x}) \ll c^2,$$

the gravitational field is homogeneous, the gravitational acceleration $\mathbf{a} = -\mathbf{g} = -\mathbf{g}(t)$ being a function of t which is uniquely determined by the motion of the origin $\mathbf{x} = 0$ relative to I.

The rate of a standard clock at rest at the point $\mathbf{x}$ as given by (100) and (132) is

$$d\tau_0 = dt(1+(\mathbf{g}\,.\,\mathbf{x})/c^2). \tag{134}$$

At the origin $\mathbf{x} = 0$ we have $\tau_0 = t$, i.e. the coordinate clock at this place is a standard clock.

It may be shown that the type of accelerated systems considered in this section together with the rotating system of § 90 are essentially the only possible rigid systems of reference in the case of non-permanent gravitational fields.

97. Rigid frames of reference moving in the direction of the X-axis

When the origin O of the system (x^ι) is moving in the direction of the X-axis of the system (X_ι), the coefficients $\alpha_{\iota k}$ are given by (IV. 154, 147), and the transformation equations (127) reduce to the equations

$$\left.\begin{aligned} X &= c\int\limits_0^t \sinh\theta\,dt+x\cosh\theta(t) \\ Y &= y, \qquad Z = z, \\ T &= \int\limits_0^t \cosh\theta\,dt+x\frac{\sinh\theta(t)}{c} \end{aligned}\right\}, \tag{135}$$

which are also obtained from (IV. 155) by the substitution (126). For the vector (130) we get, by means of (IV. 154, 146),

$$\mathfrak{g} = (g, 0, 0), \qquad g(t) = c\frac{d\theta}{dt}. \tag{136}$$

Hence, from (129),

$$ds^2 = dx^2+dy^2+dz^2-c^2\,dt^2(1+gx/c^2)^2. \tag{137}$$

This is also easily obtained directly by differentiation of (135) and substitution in the expression for ds^2 in terms of the differentials (dX_ι). In this case the gravitational field is parallel to the x-axis. The conditions (52) are satisfied everywhere except on the plane $x = -c^2/g$, where g_{44} becomes equal to zero.

In particular, if the motion of the origin O is a hyperbolic motion with the constant rest acceleration g, we have, according to (IV. 159),

$$\left.\begin{aligned} \theta(t) &= \frac{gt}{c} \\ \text{i.e.} \qquad g(t) &= c\frac{d\theta}{dt} = g = \text{constant} \end{aligned}\right\}. \tag{138}$$

In this case the gravitational field is static, the gravitational potential being

$$\chi = -\frac{c^2}{2}(g_{44}+1) = gx\left(1+\frac{gx}{2c^2}\right). \tag{139}$$

The transformation equations then reduce to

$$\left.\begin{aligned} X &= \frac{c^2}{g}\left(\cosh\frac{gt}{c}-1\right)+x\cosh\frac{gt}{c} \\ Y &= y, \qquad Z = z, \\ T &= \frac{c}{g}\sinh\frac{gt}{c}+x\frac{\sinh(gt/c)}{c} \end{aligned}\right\}, \tag{140}$$

as seen from (135) and (138) or from (IV. 160) and (126)

By elimination of the variable t in (140) we get

$$\left.\begin{aligned} X &= \frac{c^2}{g}\{[(1+gx/c^2)^2+g^2T^2/c^2]^{\frac{1}{2}}-1\} \\ Y &= y, \qquad Z = z \end{aligned}\right\}. \tag{141}$$

Thus we see that, relative to I, each reference point with constant values for x, y, z performs a hyperbolic motion in the direction of the X-axis starting at the point $X = x$, $Y = y$, $Z = z$ at $T = 0$, with zero velocity The acceleration in this hyperbolic motion is

$$\gamma = \frac{g}{1+gx/c^2} \tag{142}$$

(see § 29), and the velocity at the time T is

$$v = \frac{dX}{dT} = \frac{gT}{[(1+gx/c^2)^2+g^2T^2/c^2]^{\frac{1}{2}}} = c\tanh\frac{gt}{c} \tag{143}$$

on account of (140).

The velocity of the reference points relative to I thus depends on x, and from the point of view of an observer in I the system of reference R corresponding to the coordinates (x^i) will not appear as rigid The distance between two reference points (x, y, z) and $(x+dx, y, z)$ measured by an observer in I is found from (141) and (143).

$$dX = \frac{1+gx/c^2}{[(1+gx/c^2)^2+g^2T^2/c^2]^{\frac{1}{2}}}dx = \frac{1}{\cosh(gt/c)}dx = \sqrt{(1-v^2/c^2)}\,dx, \tag{144}$$

where $v = v(T)$ is the velocity relative to I of the system of reference R at the place considered. From the point of view of an observer in I, each part of the system R is thus contracting in accordance with the Lorentz formula.

For small values of t and x, where we can neglect terms in gx/c^2 and gt/c of higher than first order, the system of reference R considered here is identical with the system considered at the end of § 95.

Let us now consider the motion of a free particle in the gravitational field of the system (x^i). The time track of such a particle is determined by (86) with g_{ik} given by (137), and for $i = 1, 2, 3$ (86) becomes

$$\frac{d^2x}{d\tau^2} = -\left(1+\frac{gx}{c^2}\right)g\left(\frac{dt}{d\tau}\right)^2, \qquad \frac{d^2y}{d\tau^2} = \frac{d^2z}{d\tau^2} = 0. \tag{145}$$

If the particle starts from a point on the x-axis with zero velocity, the last equations (145) give

$$y = z = 0.$$

From (99) and (139) we get the following connexion between the proper time of the particle and the time variable t of the system (x^i):

$$d\tau = dt\sqrt{[(1+gx/c^2)^2-u^2/c^2]}, \tag{146}$$

where

$$u = \left|\frac{dx}{dt}\right|$$

is the velocity of the particle. (146) is equivalent to the last equation (86).

Using t instead of τ as independent variable in (145), we get by means of (146)

$$\frac{d^2x}{dt^2}-\frac{2g/c^2}{1+gx/c^2}\left(\frac{dx}{dt}\right)^2+g(1+gx/c^2) = 0. \tag{147}$$

The solution of this equation corresponding to the initial conditions

$$x = x_0, \quad \frac{dx}{dt} = 0 \quad \text{for } t = 0 \tag{148}$$

is

$$x = \frac{c^2}{g}\left[\left(1+\frac{gx_0}{c^2}\right)\frac{1}{\cosh gt/c}-1\right], \tag{149}$$

as is easily verified by differentiation and substitution in (147).

The velocity of the particle at the time t is

$$\frac{dx}{dt} = -c\left(1+\frac{gx_0}{c^2}\right)\frac{\sinh gt/c}{\cosh^2 gt/c}. \tag{150}$$

Thus with increasing t the velocity increases, reaches a maximum

$$u_{\max} = \frac{c(1+gx_0/c^2)}{2},$$

and decreases again to zero for $t \to \infty$. If $x_0 > c^2/g$, the velocity of the particle assumes values which are larger than c; however, u is always smaller than the velocity of light, which is

$$w = c(1+gx/c^2)$$

on account of (70) and (137).

For $t \to \infty$ the particle approaches the singular wall $x = -c^2/g$ of our system of coordinates. At this place also the velocity of light tends to zero, and no signals of any kind will ever reach the boundary plane.

Using the expressions (149) and (150) for x and u in (146) we get by integration for the proper time τ of the particle at the coordinate time t

$$\tau = \left(1+\frac{gx_0}{c^2}\right)\int_0^t \frac{dt}{\cosh^2 gt/c} = \left(\frac{c}{g}+\frac{x_0}{c}\right)\tanh\frac{gt}{c}. \tag{151}$$

Finally, if the origin of the system (x^i) is moving with constant velocity v in the direction of the positive X-axis, $f_\iota = 0$, and the vector g_ι defined by (130) is zero. Thus the gravitational field is zero as it should be, since the system of reference R in this case is a system of inertia moving with the velocity v. Further, since U_ι is constant and $f_\iota = U_\iota \tau$, the transformation (127) is identical with the special Lorentz transformation in this case.

98. The clock paradox

We are now in a position to state the complete solution of the clock paradox which was mentioned in § 20 and which played a certain part in the early discussions on the consistency of the theory of relativity.† Consider two standard clocks C_1 and C_2 originally situated at rest at the origin O_1 of a system of inertia S_1 with the space-time coordinates (X, Y, Z, T) (Fig 18). At the time $T = 0$ the clock C_2 is accelerated by a constant force F in the direction of the positive X-axis. When C_2 has reached the point A it has attained a certain velocity v, and from this point on C_2 is allowed to continue in a uniform motion with the constant velocity v until it reaches the point B, where it meets a constant counter-force of the same magnitude F as before, but with opposite direction. C_2 is brought to rest at C and accelerated back to B, having then attained the velocity $-v$. Between B and A it moves with the constant velocity $-v$, and at A it is attacked again by the constant force F which brings it to rest at O_1. Let $\Delta' T$, $\Delta'' T$, $\Delta''' T$ be the times which C_2 takes to travel the distances OA, AB, and BC, respectively. For symmetry reasons the motion on the way back from C to A must be just the reverse of the motion from A to C, and further we must have

$$\Delta''' T = \Delta' T.$$

† A. Einstein, *Ann. d. Phys.* **17**, 891 (1905); P. Langevin, *Scientia*, **10**, 31 (1911); M. v. Laue, *Phys. ZS.* **13**, 118 (1912), H. A Lorentz, *Das Relativitatsprinzip*, 3 Haarlemer Vorlesungen, pp. 31 and 47, Leipzig, 1914, A. Einstein, *Naturw* **6**, 697 (1918), C. Møller, *Dan. Mat. Fys. Medd.* **20**, No. 19 (1943)

Let $\Delta\tau_1$ and $\Delta\tau_2$ denote the measurements on the clocks C_1 and C_2 of the time elapsed between the two encounters of the clocks, τ_1 and τ_2 being the proper times of C_1 and C_2, respectively. Since C_1 is constantly at rest at the point O, $\Delta\tau_1$ is equal to the total increase ΔT in the time variable T of the system S_1 between the two encounters. Thus we have

$$\Delta\tau_1 = \Delta T = 2(\Delta' T+\Delta'' T+\Delta''' T) = 2(2\Delta' T+\Delta'' T). \qquad (152)$$

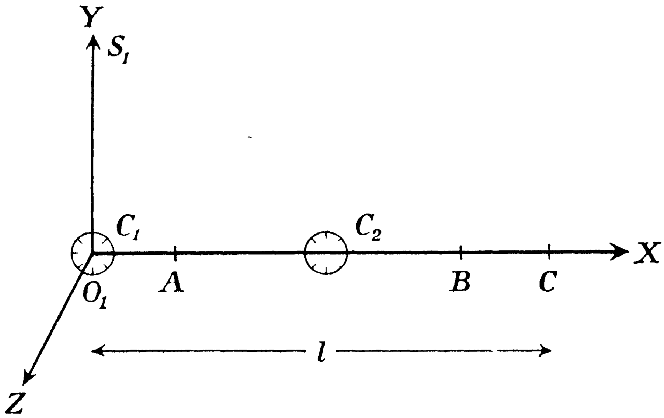

FIG 18

Similarly we get

$$\Delta\tau_2 = 2(\tau_2'+\tau_2''+\tau_2''') = 2(2\tau_2'+\tau_2''), \qquad (153)$$

where τ_2', τ_2'', τ_2''' denote the increase in proper time of C_2 during its travel through OA, AB, and BC, respectively The motion of the clock C_2 from O to A is a hyperbolic motion and is thus described by the equation (III. 47), i e.

$$X = \frac{c^2}{g}\left\{\left[1+\left(\frac{gT}{c}\right)^2\right]^{\frac{1}{2}}-1\right\}, \qquad (154)$$

where
$$g = \frac{F}{m_0}$$

and m_0 is the rest mass of C_2. The velocity $u = dX/dT$ is thus

$$u = \frac{dX}{dT} = \frac{gT}{\sqrt{[1+(gT/c)^2]}}. \qquad (155)$$

Hence we get
$$v = \frac{g\,\Delta' T}{\sqrt{[1+(g\Delta' T/c)^2]}} \qquad (156)$$

or
$$g\,\Delta' T = \frac{v}{\sqrt{(1-v^2/c^2)}}. \qquad (157)$$

Using (155) we can now calculate the proper time τ_2' of C_2 by means of the formula (II. 38) of the special theory of relativity valid in the system of inertia S_1. Hence

$$\tau_2' = \int\limits_0^{\Delta' T} \sqrt{(1-u^2/c^2)}\, dT = \int\limits_0^{\Delta' T} \frac{dT}{\sqrt{[1+(gT/c)^2]}} = \frac{c}{g}\sinh^{-1}\frac{g\,\Delta' T}{c}.$$

On account of (157) this may also be written

$$\frac{g\,\Delta' T}{c} = \frac{v/c}{\sqrt{(1-v^2/c^2)}} = \sinh\frac{g\tau_2'}{c} = \sinh\frac{g\tau_2'''}{c} \tag{158}$$

or

$$\tanh\frac{g\tau_2'}{c} = \frac{\sinh\frac{g\tau_2'}{c}}{\sqrt{\left(1+\sinh^2\frac{g\tau_2'}{c}\right)}} = \frac{v}{c}. \tag{158'}$$

In the same way we get from (II. 38)

$$\tau_2'' = \Delta'' T\sqrt{(1-v^2/c^2)}. \tag{159}$$

Now, if for constant value of v we apply a larger and larger force F the acceleration $g = F/m_0$ will also increase. In the limit $g \to \infty$ for constant v we see from (158) that both $\Delta' T = \Delta''' T$ and $\tau_2' = \tau_2'''$ tend to zero. In this limit, where the velocity v is attained nearly instantaneously, we thus get from (152), (153), and (159)

$$\Delta\tau_1 = 2\Delta'' T, \qquad \Delta\tau_2 = 2\tau_2'' = 2\Delta'' T\sqrt{(1-v^2/c^2)}, \tag{160}$$

i.e.

$$\Delta\tau_2 = \Delta\tau_1\sqrt{(1-v^2/c^2)} \tag{161}$$

as we should have. The moving clock C_2 is lagging behind the stationary clock C_1. Further, in the limit $g \to \infty$, the maximum distance l between the two clocks is simply given by

$$l = v\,\Delta'' T. \tag{162}$$

We shall now see that the same result is obtained if the whole process is treated in a rigid system of reference S_2 with coordinates (x, y, z, t) which follows the motion of C_2 in such a way that C_2 is permanently situated at the origin. In the time intervals where S_2 is accelerated relative to S_1 or to the distant stars, we have a gravitational field in S_2. During the time interval $0 < t < \tau_2'$ of magnitude $\Delta' t = \tau_2'$, the gravitational field is described by the scalar potential (139). In the interval $\tau_2' < t < \tau_2'+\tau_2''$ of magnitude $\Delta'' t = \tau_2''$, we have $\chi = 0$ and in the interval $\tau_2'+\tau_2'' < t < \tau_2'+\tau_2''+\tau_2'''$ of magnitude $\Delta''' t = \tau_2''' = \tau_2' = \Delta' t$ we have $\chi = -gx(1-gx/2c^2)$. During the first period $\Delta' t$ the clock C_1 is falling freely in the direction of the negative x-axis in accordance with the equation of motion (149). In the period $\Delta''t$ it is moving uniformly with the velocity $-v$ and, finally, in the period $\Delta'''t$ it is brought to rest at a point with the coordinates $x_0 = -l$. Since the systems S_1 and S_2 are at rest relative to each other at that moment, the maximum distance between the two clocks is the same when measured in S_2 or in S_1. After this moment the clock C_1 returns to the origin with reversed motion. The clock C_2 is at rest at the origin of S_2 during the whole process, since the gravitational field is counterbalanced by the external force F.

The increase in the proper time of the clock C_1 can now be calculated by means of the general formula (99) with the expressions for χ given above. In (149), (150), and (151) we have given the solution of the equations of motion (147). If τ_1', τ_1'', and τ_1''' denote the increase of the proper time of C_1 in the intervals $\Delta't$, $\Delta''t$, $\Delta'''t$, respectively, we have obviously

$$\Delta\tau_1 = 2(\tau_1' + \tau_1'' + \tau_1'''). \tag{163}$$

Similarly, since C_2 is at rest at the origin $x = 0$, where χ is constantly equal to zero, we have

$$\Delta\tau_2 = 2(\Delta't + \Delta''t + \Delta'''t) = 2(2\Delta't + \Delta''t) = 2(2\tau_2' + \tau_2''). \tag{164}$$

Since C_1 starts from the origin with zero velocity, we can get τ_1' from (151) by putting $x_0 = 0$ and $t = \Delta't = \tau_2'$. Hence

$$\tau_1' = \frac{c}{g}\tanh\frac{g\tau_2'}{c} = \frac{v}{g} \tag{165}$$

on account of (158').

During the interval $\Delta''t$, the clock C_1 is moving with constant velocity in a field-free space. Thus we get

$$\tau_1'' = \Delta''t\sqrt{(1 - v^2/c^2)} = \tau_2''\sqrt{(1 - v^2/c^2)} = \Delta''T(1 - v^2/c^2) \tag{166}$$

on account of (159). Finally, τ_1''' may be obtained by putting g equal to $-g$, $x_0 = -l$, and $t = \Delta'''t = \tau_2''' = \tau_2'$ in (151). Hence

$$\tau_1''' = \left(\frac{c}{g} + \frac{l}{c}\right)\tanh\frac{g\tau_2'}{c} = \left(\frac{c}{g} + \frac{l}{c}\right)\frac{v}{c} \tag{167}$$

on account of (158').

In the limit as $g \to \infty$ with constant v we get from (165) $\tau_1' \to 0$. But although $\Delta'''t = \tau_2''' = \tau_2' \to 0$ in this limit, τ_1''' approaches the finite limit

$$\tau_1''' = \frac{lv}{c^2}. \tag{168}$$

This surprising result is due to the influence on the rate of the clock C_1 of the gravitational scalar potential $\chi = -gx(1 - gx/2c^2)$ which becomes infinite in the limit $g \to \infty$.

Thus, in the limit $g \to \infty$, we get from (163), (166), (168), (162), (164), and (159),

$$\left.\begin{aligned}\Delta\tau_1 &= 2\left[\Delta''T\left(1-\frac{v^2}{c^2}\right)+\frac{\Delta''Tv^2}{c^2}\right] = 2\Delta''T\\ \Delta\tau_2 &= 2\tau_2'' = 2\Delta''T\sqrt{(1-v^2/c^2)}\end{aligned}\right\}, \tag{169}$$

i.e. the same result as in (160). This result which represents the solution of the clock paradox is, of course, not surprising, since the proper time is an invariant which has the same value in any system of coordinates.

We shall now finally consider another much simpler example of the same phenomenon in which also the importance of the gravitational vector potential for the rate of moving clocks is illustrated. Consider a clock C_2 which under the influence of a central force F performs a uniform circular motion in a system of inertia S_1. If the radius of the circle is R and the constant angular velocity ω_1, the velocity of the particle is $R\omega_1$ and the increase in the proper time τ_2 during a revolution is

$$\tau_2 = T\sqrt{(1-R^2\omega^2/c^2)} = \frac{2\pi}{\omega}\sqrt{(1-R^2\omega^2/c^2)} \tag{170}$$

according to the formula (II. 38) The corresponding increase in the proper time of a clock C_1 at rest at the periphery of the circle is

$$\tau_1 = T = \frac{2\pi}{\omega}. \tag{171}$$

Let us now treat the same phenomenon from the point of view of an observer on the rotating disk of § 90. In this system S_2 the clock C_2 is at rest at the point ($r = R$, $\vartheta = 0$), say, while the clock C_1 is rotating with the angular velocity

$$\frac{d\vartheta}{dt} = -\omega$$

in a circle of radius $r = R$. C_1 is falling freely under the influence of the gravitational field with the potentials (74) and (97),

$$\chi = -\tfrac{1}{2}r^2\omega^2, \qquad \gamma_i = \left(0, \frac{\omega r^2}{c\sqrt{(1-r^2\omega^2/c^2)}}, 0\right). \tag{172}$$

It is easily seen that r = constant, $d\vartheta/dt = -\omega$ is a solution of the equations of motion (86) with g_{ik} given by (73). The clock C_2 remains at rest, the gravitational acceleration (96) being counter-balanced by the force F.

For the increase τ_2 of the proper time of C_2 during the time $t = 2\pi/\omega$ we now get at once from the formula (100) for a clock at rest

$$\tau_2 = \frac{2\pi}{\omega}\sqrt{(1+2\chi/c^2)} = \frac{2\pi}{\omega}\sqrt{(1-r^2\omega^2/c^2)} \tag{173}$$

in accordance with (170).

To determine the corresponding increase of the proper time of C_1 we have to use the general formula (98). Since

$$u^\iota = \frac{dx^\iota}{dt} = (0, -\omega, 0),$$

we get from (74)

$$u^2 = \frac{d\sigma^2}{dt^2} = \gamma_{\iota\kappa}\frac{dx^\iota}{dt}\frac{dx^\iota}{dt} = \gamma_{22}\,\omega^2 = \frac{r^2\omega^2}{1-r^2\omega^2/c^2}.$$

Further, using (172),

$$\gamma_\iota u^\iota = -\frac{r^2\omega^2}{c\sqrt{(1-r^2\omega^2/c^2)}}.$$

Hence, from (98),

$$\tau_1 = \frac{2\pi}{\omega}\left\{\left(\sqrt{(1-r^2\omega^2/c^2)}+\frac{r^2\omega^2}{c^2\sqrt{(1-r^2\omega^2/c^2)}}\right)^2-\frac{r^2\omega^2}{c^2(1-r^2\omega^2/c^2)}\right\}^{\frac{1}{2}} = \frac{2\pi}{\omega}. \tag{174}$$

Thus, in this case, the effects of the gravitational potentials and of the velocity of the clock C_1 on the proper time τ_1 as expressed in the formula (98) just cancel and for τ_1 and τ_2 we get the same values as before. The physical interpretation of the effect is, however, completely different in the two cases In S_1 the effect is ascribed to the velocity of the particles only, while in S_2 the phenomenon is explained as a joint effect of the gravitational field and the motion.

IX

PERMANENT GRAVITATIONAL FIELDS. TENSOR CALCULUS IN A GENERAL RIEMANNIAN SPACE

99. Four-dimensional formulation of the general principle of relativity and of the principle of equivalence

In the preceding chapter we have considered the case of gravitational fields which could be transformed away by the introduction of the pseudo-Cartesian coordinates of the system of inertia I from which we started in § 88. We have seen that the action of the gravitational field in an arbitrary system of coordinates (x^i) is described by the metric tensor g_{ik} which determines the line element in space-time by the equation

$$ds^2 = g_{ik}\, dx^i dx^k. \tag{1}$$

The non-permanent gravitational fields are thus characterized by the property that the interval can be brought into the form (VIII. 40)

$$ds^2 = G_{ik}\, dX^i dX^k \tag{2}$$

for all points in 4-space by a suitable choice of the space-time coordinates or, in other words, the space is a flat pseudo-Euclidean space in this case.

According to the principle of equivalence there should, however, be no essential difference between permanent and non-permanent fields, both types of fields satisfying the same fundamental laws. We shall therefore assume that the gravitational fields produced by the presence of large masses as, for instance, that of the earth or the sun, are described by the metric tensor g_{ik} in 4-space in the same way as in the case of the artificially produced non-permanent fields. In particular, it is assumed that the time tracks of a free (i.e. freely falling) particle and of a light ray traversing permanent gravitational fields are geodesic lines in 4-space, given by the same equations (VIII. 86) and (VIII. 87) as in the case of non-permanent fields. The only difference will then be that the permanent fields cannot be removed completely by a suitable transformation of the space-time coordinates, i.e. ds^2 cannot be brought into the form (2) simultaneously for all points in 4-space. Hence, in this case, the 4-space is a 'curved' space with a general Riemannian geometry.

As we shall see in § 104, it is, however, always possible in an infinite number of ways to introduce a so-called geodesic system of coordinates $(\mathring{x}^i)$ for which the first derivatives $\partial \mathring{g}_{ik}/\partial \mathring{x}^l$ of the metric tensor are zero and the values of $\mathfrak{g}_{ik}$ are equal to G_{ik} at a given point in 4-space. Geometrically this means that the space may be treated as flat in an infinitesimal region around every point, in analogy to the two-dimensional case where a curved surface may be replaced by the tangential plane in a small region around the point considered. Systems of space-time coordinates ($\mathfrak{x}^i$) with the above-mentioned properties may also be called 'local systems of inertia'; for in the case of permanent gravitational fields they play locally the same role as the systems of inertia in the case of non-permanent fields. For an arbitrary choice of space-time coordinates (x^ι), i.e. for an arbitrary numbering of the events in physical space, the quantities $g_{\iota k}$ can therefore always be experimentally determined by the same methods as those described in § 89.

According to the general principle of relativity, the laws of nature must now be expressible in the form of equations which are form-invariant. Thus, if the law in question is expressed by equations of the form

$$F\left(A, B, \ldots; \frac{\partial A}{\partial x^\iota}, \frac{\partial B}{\partial x^\iota}, \ldots\right) = 0, \tag{3}$$

where A, B,... are physical quantities, we have in another arbitrary system of coordinates (x'^ι) the same functional relation between the physical quantities in (x'^ι), i.e.

$$F\left(A', B', \ldots; \frac{\partial A'}{\partial x'^\iota}, \frac{\partial B'}{\partial x'^\iota}, \ldots\right) = 0. \tag{3'}$$

The only difference from the case considered in § 35 is that now the gravitational quantities $g_{\iota k}$ will be among the set of physical quantities A, B,... entering into the equations (3).

In the special theory of relativity, the covariance of the laws of nature under Lorentz transformations could be very elegantly expressed by means of the four-dimensional tensor calculus. To obtain a similar representation of the laws of nature in the general theory of relativity we shall have to make a generalization of the tensor calculus developed in Chapter IV for pseudo-Cartesian systems of coordinates.

In the case of non-permanent gravitational fields, this generalization simply consists in the rather trivial extension of the notion of vectors and tensors to general curvilinear space-time coordinates, the geometry of 4-space being the same as in the special theory of relativity. In the general case of permanent fields, however, the structure of the 4-space itself is different and the development of a tensor calculus for a general Riemannian space will be required. Formally the difference between these two cases is not very great and, as we shall see, most of the tensor relations holding for curvilinear coordinates in a flat space can be used also in a general curved space.

100. Contravariant and covariant components of a four-vector

Let (x^i) be an arbitrary system of curvilinear coordinates in 4-space. The geometry in this space is then completely defined if the components g_{ik} of the metric tensor are given functions of the space-time coordinates. We shall assume that the g_{ik} satisfy the conditions (VIII. 52) at every event point This means that the determinant

$$g = |g_{ik}| = \begin{vmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{vmatrix} \tag{4}$$

is negative. Let A_{ik} be the conjugate minor of the element g_{ik} in the ith row and kth column. From well-known theorems of the theory of determinants we then get

$$\left.\begin{aligned} g &= \sum_k g_{ik} A_{ik} \\ &\text{(no summation over } i) \\ \text{and} \qquad \sum_l g_{il} A_{kl} &= 0 \quad \text{for} \quad i \neq k \end{aligned}\right\}. \tag{5}$$

Defining now a symmetrical scheme of quantities g^{ik} by

$$g^{ik} = \frac{A_{ik}}{g} = g^{ki}, \tag{6}$$

we have, according to (5),

$$g_{il} g^{kl} = g_{li} g^{lk} = \delta_i^k, \tag{7}$$

where

$$\delta_i^k = \begin{cases} 1 \text{ for } i = k \\ 0 \text{ for } i \neq k \end{cases} \tag{8}$$

If we introduce a new system of coordinates by the transformation

$$x'^i = x'^i(x^l) \tag{9}$$

we get, exactly as in the case of a flat space considered in § 88 (see equations VIII. 54-58),

$$\left.\begin{aligned} dx'^i &= \alpha_k^i\, dx^k = \frac{\partial x'^i}{\partial x^k}\, dx^k \\ dx^i &= \breve{\alpha}_k^i\, dx'^k = \frac{\partial x^i}{\partial x'^k}\, dx'^k \end{aligned}\right\} \tag{10}$$

$$\alpha_l^i \breve{\alpha}_k^l = \breve{\alpha}_l^i \alpha_k^l = \delta_k^i \tag{11}$$

$$\left.\begin{aligned} g'_{ik} &= \breve{\alpha}_i^l \breve{\alpha}_k^m g_{lm} \\ g_{ik} &= \alpha_i^l \alpha_k^m g'_{lm} \end{aligned}\right\}. \tag{12}$$

In virtue of the transformations (10), (12) the expression for the interval is invariant

$$ds^2 = g_{ik}\, dx^i dx^k = g'_{ik}\, dx'^i dx'^k. \tag{13}$$

It is always possible to choose the transformation (9) so as to make g'_{ik} equal to the G_{ik} defined by (VIII. 41) at a given point in 4-space. In general it will, however, not be possible to obtain this result simultaneously for all points; for the coefficients α^i_k cannot be chosen freely since they must satisfy the integrability conditions

$$\frac{\partial \alpha^i_k}{\partial x^l} = \frac{\partial \alpha^i_l}{\partial x^k} \tag{14}$$

following from the definition of the α^i_k in (10). This will be possible only for a flat space, in which case the functions g_{ik} satisfy special conditions (see § 107).

A vector at a definite point (x^i) is now defined as a quantity which has four components a^i in every system of coordinates satisfying the same transformation law as the coordinate differentials in (10), i.e.

$$a'^i = \alpha^i_k a^k. \tag{15}$$

On account of (11) the reciprocal relations are

$$a^i = \breve{\alpha}^i_k a'^k. \tag{15'}$$

Since the coefficients α^i_k and $\breve{\alpha}^i_k$ are functions of the coordinates (x^i), the transformation laws (15) and (15′) have a definite meaning only with reference to the particular point with which the vector a^i is connected.

In curvilinear systems of coordinates we have to distinguish between the *contravariant* components with the transformation equations (15) and the *covariant* components a_i of the *same* vector, defined by the equations

$$a_i = g_{ik} a^k \tag{16}$$

in every system of coordinates. On account of (7) the relations reciprocal to (16) are

$$a^i = g^{ik} a_k. \tag{17}$$

The operations (16) and (17) are called *lowering* and *raising* of indices. In a Cartesian system of coordinates in Euclidean space we have

$$g_{ik} = \delta_{ik}$$

and there will be no difference between the covariant and the contravariant components.

From (12), (15), and (11) we get the following transformation equations for the covariant components:

$$a'_\iota = g'_{\iota k} a'^k = \breve{\alpha}^l_\iota \breve{\alpha}^m_k \alpha^k_n g_{lm} a^n = \breve{\alpha}^l_\iota \delta^m_n g_{lm} a^n = \breve{\alpha}^l_\iota g_{lm} a^m$$

or
$$a'_\iota = \breve{\alpha}^k_\iota a_k, \tag{18}$$
the reciprocals of which are
$$a_\iota = \alpha^k_\iota a'_k. \tag{19}$$

In a flat (3+1)-space we can use pseudo-Cartesian coordinates and the connexion between the covariant and the contravariant components of a vector in this system of coordinates is simply

$$\left.\begin{aligned} a_\iota &= G_{\iota k} a^k = (a^1, a^2, a^3, -a^4) \\ \text{or}\quad a^\iota &= (a_1, a_2, a_3, -a_4) = G^{\iota k} a_k \end{aligned}\right\}, \tag{20}$$

i.e. in this case we have
$$G^{\iota k} = G_{\iota k} \tag{20'}$$

in accordance with (7) and (VIII. 41). If we had used the real representation of vectors in the special theory of relativity, it would have been necessary to distinguish between covariant and contravariant components in Chapter IV. It was just in order to avoid this slight complication that the imaginary time components were introduced by which the 4-space was made formally Euclidean. In any case, it is easy to pass over from the imaginary to the real representation. It is only to be remembered that the contravariant components of a vector in the real representation are obtained from the components in the imaginary representation by dropping the symbol i in the fourth component. The covariant components of the vector are then obtained from (20).

From this rule it follows that the norm of a vector, which in the imaginary representation was defined by (IV. 25), in the real representation takes the form
$$a^2 = a_\iota a^\iota, \tag{21}$$

and, since this expression is invariant for all coordinate transformations, (21) will hold also for curvilinear coordinates. In fact, we get from (18), (15), and (11)

$$a'_i a'^\iota = \breve{\alpha}^l_\iota \alpha^\iota_m a_l a^m = \delta^l_m a_l a^m = a_l a^l. \tag{22}$$

We may therefore take (21) as the norm of a vector also in a general Riemannian space. According to (16) and (17), (21) may be written in the different forms
$$a^2 = a_\iota a^i = g_{\iota k} a^i a^k = g^{\iota k} a_\iota a_k. \tag{23}$$

Similarly, the vector product of two vectors a and b is given by the invariant
$$a_i b^i = g_{\iota k} a^i b^k = a^\iota b_\iota = g^{\iota k} a_\iota b_k. \tag{24}$$

101. Tensor algebra

The generalization of the tensor calculus developed in §§ 39–44 for Cartesian systems of coordinates to the general curvilinear coordinates of Riemannian space is now obvious. A tensor of rank n in 4-space is a quantity with 4^n components which transform with respect to each index like a vector, i.e. the transformation is given by (15) or (18), according as the tensor is contravariant or covariant with respect to the index in question. The connexions between the covariant and contravariant components of a tensor are given by the general rules (16) and (17) for lowering and raising indices.

The transformation laws for the contravariant and covariant components of a tensor of rank 2 are thus

$$t'^{ik} = \alpha^i_l \alpha^k_m t^{lm} \tag{25}$$

$$t'_{ik} = \breve{\alpha}^l_i \breve{\alpha}^m_k t_{lm}, \tag{26}$$

respectively, and the connexions between the covariant and contravariant components are

$$\left.\begin{aligned} t_{ik} &= g_{il} g_{km} t^{lm} \\ t^{ik} &= g^{il} g^{km} t_{lm} \end{aligned}\right\} \tag{27}$$

in every system of coordinates. On account of (11) and (12), the equations (25–27) are easily seen to be compatible. Besides the purely contravariant and covariant components we can also form the mixed components

$$t_{i\cdot}^{\ k} = g^{kl} t_{il} = g_{il} t^{lk} \quad \text{and} \quad t^i_{\ \cdot k} = g^{il} t_{lk} = g_{kl} t^{il}, \tag{28}$$

which transform according to the laws

$$t'^{\ k}_{i\cdot} = \breve{\alpha}^l_i \alpha^k_m t_l^{\ m}, \qquad t'^i_{\ \cdot k} = \alpha^i_l \breve{\alpha}^m_k t^l_{\ m}. \tag{29}$$

In general, $t^i_{\ k}$ will be different from $t_k^{\ i}$.

All symmetry properties like

$$t^{ik} = \pm t^{ki} \tag{30}$$

are invariant properties. They may also be expressed as

$$t_{ik} = \pm t_{ki}, \qquad t_{i\cdot}^{\ k} = \pm t^k_{\ i} \tag{31}$$

which, on account of (27), (28), are equivalent to (30). Thus the mixed components $t_{i\cdot}^{\ k}$ and $t^k_{\ i}$ of a symmetric tensor are equal and may therefore simply be written as t^k_i.

A comparison of (12) and (26) shows that the quantities g_{ik} themselves are the covariant components of a symmetrical tensor of rank 2—the metric tensor. Further, we see from (7) that the mixed components g^k_i of this tensor are given by the Kronecker symbol δ_i^k and that the contravariant components of the metric tensor are equal to the quantities g^{ik} defined by (6).

Similarly, as in the case of Cartesian coordinates considered in § 41, we can now also in the general case form new tensors by the processes of addition, direct multiplication, and contraction. By addition of two tensors of rank n we get a new tensor of rank n, and by direct multiplication of two tensors of the ranks n and m we get a tensor of rank $n+m$. It should be remarked, however, that these processes have an unambiguous meaning in the general case only if the two tensors belong to the same point in 4-space. Finally, the process of contraction which in curvilinear systems of coordinates consists in equating an upper and a lower index, and summing, reduces the rank of a tensor by 2. By contraction of a tensor of rank 2 we thus get a tensor of rank 0, i.e. an invariant. From the transformation law (29) we see at once that

$$t'^{\,i}_{i} = \breve{\alpha}^l_i \alpha^i_m t_l^{\,m} = \delta^l_m t_l^{\,m} = t_l^{\,l} \tag{32}$$

is an invariant which, by means of (28), may be written in the different forms

$$t_i^{\,i} = g_{ik} t^{ki} = g^{ik} t_{ik} = t^k{}_k. \tag{33}$$

An example of a combined application of the processes of direct multiplication and contraction is offered by the equation (24).

102. Pseudo-tensors. Dual tensors

Let $\alpha = |\alpha_i^k|$ and $\breve{\alpha} = |\breve{\alpha}_i^k|$ be the determinants corresponding to the scheme of transformation coefficients α_i^k and $\breve{\alpha}_i^k$, respectively. According to the multiplication rules for determinants, we get at once from (11)

$$\alpha . \breve{\alpha} = 1, \tag{34}$$

and from (12) $$g' = |g'_{ik}| = \breve{\alpha} . g . \breve{\alpha} = \breve{\alpha}^2 g = \frac{g}{\alpha^2},$$

i.e. $$\sqrt{(-g')} = |\breve{\alpha}| \sqrt{(-g)} = \frac{1}{|\alpha|} \sqrt{(-g)}, \tag{35}$$

where $|\breve{\alpha}|$ and $|\alpha|$ are the absolute values of the determinants $\breve{\alpha}$ and α.

A pseudo-tensor is now defined as a quantity whose components transform like the components of a tensor, except that they are multiplied by the sign $\frac{\alpha}{|\alpha|} = \frac{\breve{\alpha}}{|\breve{\alpha}|}$ of the transformation determinants. If α and consequently $\breve{\alpha}$ are positive, a pseudo-tensor therefore transforms like a tensor of the same rank. In the same way as in § 43, it is now easily seen that a quantity A_{iklm}, whose components in every system of coordinates are equal to the Levi-Civita symbol δ_{iklm}, transforms according to the law

$$A'_{iklm} = \alpha \breve{\alpha}^r_i \breve{\alpha}^s_k \breve{\alpha}^t_l \breve{\alpha}^u_m A_{rstu}. \tag{36}$$

From (35) and (36) we thus see that the quantities

$$\epsilon_{iklm} = \sqrt{(-g)}\, \delta_{iklm} \tag{37}$$

are the covariant components of a completely antisymmetric pseudo-tensor of rank 4.

In the same way as in § 44, we can now to an antisymmetrical tensor of rank n adjoin a dual pseudo-tensor of rank $4-n$ by means of the pseudo-tensor (37). Thus, if F^{ik} are the contravariant components of an antisymmetrical tensor, the covariant components of the dual tensor are given by

$$F^*_{ik} = \tfrac{1}{2}\epsilon_{iklm} F^{lm} = \tfrac{1}{2}\sqrt{(-g)}\, \delta_{iklm} F^{lm}, \tag{38}$$

i.e.

$$\left.\begin{aligned} F^*_{23} &= \sqrt{(-g)}F^{14}, & F^*_{31} &= \sqrt{(-g)}F^{24}, & F^*_{12} &= \sqrt{(-g)}F^{34} \\ F^*_{14} &= \sqrt{(-g)}F^{23}, & F^*_{24} &= \sqrt{(-g)}F^{31}, & F^*_{34} &= \sqrt{(-g)}F^{12} \end{aligned}\right\}. \tag{39}$$

Two infinitesimal vectors a^i, b^k define a parallelogram described by an antisymmetrical tensor with the contravariant components

$$\sigma^{ik} = a^i b^k - a^k b^i, \tag{40}$$

the area σ being given by $\sigma^2 = \frac{1}{2}\sigma_{ik}\sigma^{ik}$.

The corresponding dual tensor

$$\sigma^*_{ik} = \sqrt{(-g)}\, \delta_{iklm} a^l b^m \tag{41}$$

is orthogonal to σ^{ik}, i.e. $\sigma^*_{ik}\sigma^{ik} = 0$.

Three infinitesimal vectors a^i, b^k, c^l define a three-dimensional parallelepiped described by the antisymmetrical tensor

$$V^{ikl} = \begin{vmatrix} a^i & b^i & c^i \\ a^k & b^k & c^k \\ a^l & b^l & c^l \end{vmatrix} \tag{42}$$

or by its dual pseudo-vector

$$V_i = \frac{1}{3!}\epsilon_{iklm} V^{klm} = \frac{\sqrt{(-g)}}{3!}\delta_{iklm} V^{klm} = \sqrt{(-g)}\, \delta_{iklm} a^k b^l c^m \tag{43}$$

which is orthogonal to the vectors a^i, b^i, c^i, i.e.

$$V_i a^i = V_i b^i = V_i c^i = 0.$$

The volume V of the parallelepiped is given by

$$V^2 = -V_i V^i = -\frac{1}{3!} V_{klm} V^{klm}. \tag{44}$$

Finally, four infinitesimal vectors a^i, b^k, c^l, d^m define a four-dimensional parallelepiped described by the tensor

$$\Sigma^{iklm} = \begin{vmatrix} a^i & b^i & c^i & d^i \\ a^k & b^k & c^k & d^k \\ a^l & b^l & c^l & d^l \\ a^m & b^m & c^m & d^m \end{vmatrix} \tag{45}$$

or by the dual pseudo-invariant

$$\Sigma = \frac{1}{4!}\epsilon_{iklm}\Sigma^{iklm} = \frac{\sqrt{(-g)}\,\delta_{iklm}}{4!}\Sigma^{iklm} = \sqrt{(-g)}\,\delta_{iklm}\,a^i b^k c^l d^m,$$

i.e.

$$\Sigma = \sqrt{(-g)}\begin{vmatrix} a^1 & b^1 & c^1 & d^1 \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \\ a^4 & b^4 & c^4 & d^4 \end{vmatrix}. \tag{46}$$

If a^i, b^i, c^i, d^i are infinitesimal vectors lying in the directions of the coordinate curves and of lengths dx^1, dx^2, dx^3, dx^4, respectively, we have $a^i = (dx^1, 0, 0, 0)$, $b^i = (0, dx^2, 0, 0)$, etc., and the corresponding four-dimensional volume element is given by the pseudo-invariant

$$d\Sigma = \sqrt{(-g)}\,dx^1dx^2dx^3dx^4. \tag{47}$$

This is the generalization of the expression (VIII. 17) for the volume element in a two-dimensional space with positive definite metric. In a 3-space we can similarly adjoin a pseudo-tensor of rank $3-n$ to an antisymmetrical tensor of rank n. (See Appendix 5.)

103. Geodesic lines. Christoffel's formulae

The geodesic lines are defined by the equations (VIII. 30)

$$\frac{d}{d\lambda}\left(g_{ik}\frac{dx^k}{d\lambda}\right) = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda}, \tag{48}$$

where λ is an arbitrary invariant parameter and the indices i, k, l run from 1 to 4. (48) may also be written

$$g_{ik}\frac{d^2x^k}{d\lambda^2}+\left(\frac{\partial g_{ik}}{\partial x^l}-\frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\right)\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda} = 0.$$

Multiplying this equation by g^{mi} we get, on account of (7),

$$\frac{d^2x^m}{d\lambda^2}+g^{mi}\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^l}+\frac{\partial g_{il}}{\partial x^k}-\frac{\partial g_{kl}}{\partial x^i}\right)\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda} = 0,$$

or

$$\frac{d^2x^i}{d\lambda^2}+\Gamma^i_{kl}\frac{dx^k}{d\lambda}\frac{dx^l}{d\lambda} = 0 \tag{49}$$

with

$$\left.\begin{aligned} \Gamma^{\iota}_{kl} &= g^{\iota m}\Gamma_{m,kl} \\ \Gamma_{\iota,kl} &= \frac{1}{2}\left(\frac{\partial g_{\iota k}}{\partial x^{l}}+\frac{\partial g_{\iota l}}{\partial x^{k}}-\frac{\partial g_{kl}}{\partial x^{\iota}}\right) = g_{\iota r}\Gamma^{r}_{kl} \end{aligned}\right\}. \tag{50}$$

The quantities Γ^{ι}_{kl}, $\Gamma_{\iota,kl}$ defined by (50) are the *Christoffel three-index symbols*. They obviously satisfy the relations

$$\Gamma^{\iota}_{kl} = \Gamma^{\iota}_{lk}, \qquad \Gamma_{\iota,kl} = \Gamma_{\iota,lk}, \qquad \frac{\partial g_{\iota k}}{\partial x^{l}} = \Gamma_{\iota,kl}+\Gamma_{k,\iota l}. \tag{51}$$

The connexion between Γ^{ι}_{kl} and $\Gamma_{\iota,kl}$ is the same as that between the covariant and contravariant components of a tensor but, as we shall see, the Christoffel symbols do not transform like a tensor.

Since the equations (48) or (49) are the Euler equations corresponding to the invariant variational principle (VIII. 21, 29), they must hold in every system of coordinates. By differentiation of the equations (10) we therefore get

$$\left.\begin{aligned} \frac{dx^{\iota}}{d\lambda} &= \breve{\alpha}^{\iota}_{r}\frac{dx'^{r}}{d\lambda} \\ \frac{d^{2}x^{\iota}}{d\lambda^{2}} &= \breve{\alpha}^{\iota}_{r}\frac{d^{2}x'^{r}}{d\lambda^{2}}+\frac{\partial\breve{\alpha}^{\iota}_{r}}{\partial x'^{s}}\frac{dx'^{r}}{d\lambda}\frac{dx'^{s}}{d\lambda} \\ &= \left(-\breve{\alpha}^{\iota}_{r}\Gamma'^{r}_{kl}+\frac{\partial\breve{\alpha}^{\iota}_{k}}{\partial x'^{l}}\right)\frac{dx'^{k}}{d\lambda}\frac{dx'^{l}}{d\lambda} \end{aligned}\right\}, \tag{52}$$

where we have used the equations (49) in the system of coordinates (x'^{ι}).

Using (52) in (49) we get

$$\left(-\breve{\alpha}^{\iota}_{r}\Gamma'^{r}_{kl}+\frac{\partial\breve{\alpha}^{\iota}_{k}}{\partial x'^{l}}+\Gamma^{\iota}_{rs}\breve{\alpha}^{r}_{k}\breve{\alpha}^{s}_{l}\right)\frac{dx'^{k}}{d\lambda}\frac{dx'^{l}}{d\lambda} = 0,$$

and, since this equation must hold for independent values of the variables $dx'^{k}/d\lambda$, the expression inside the brackets, which is symmetrical in k and l, must be zero. If we multiply the relation obtained in this way by α^{m}_{ι} we get, on account of (11), the Christoffel formulae

$$\Gamma'^{\iota}_{kl} = \alpha^{\iota}_{r}\frac{\partial\breve{\alpha}^{r}_{k}}{\partial x'^{l}}+\alpha^{\iota}_{r}\breve{\alpha}^{s}_{k}\breve{\alpha}^{t}_{l}\Gamma^{r}_{st}. \tag{53}$$

By linear (affine) transformations, where the coefficients $\breve{\alpha}^{i}_{k}$, α^{ι}_{k} are constant, the first term on the right-hand side of (53) is zero and the Christoffel symbols transform like the components of a tensor. For more general transformations this will not be the case, and the Γ^{ι}_{kl}, $\Gamma_{i\,kl}$ are therefore called *affine tensors*.

If the $g_{\iota k}$ are constant, as in the case of the pseudo-Cartesian system of coordinates in a flat space, the Christoffel symbols (50) vanish.

104. Local systems of inertia

In general, it is not possible to introduce a system of coordinates which makes the components of the metric tensor independent of the coordinates, but, as we shall see now, we can always ensure that this is approximately true in the immediate surroundings of a given point P in 4-space. More exactly, we can always find a system of coordinates $(\overset{0}{x}_i)$—a *geodesic* system—for which $\partial \overset{0}{g}_{ik}/\partial \overset{0}{x}^l = 0$ at the point P. Let (x^i) be the original system of coordinates, and let $\Gamma^i_{kl}(P)$ be the values of the Christoffel symbols at the point P. If (x^i_P) denote the coordinates of this point, the transformation

$$\overset{0}{x}^i = x^i - x^i_P + \tfrac{1}{2}\Gamma^i_{rs}(P)(x^r - x^r_P)(x^s - x^s_P) \tag{54}$$

will lead to the desired result. If we identify the primed system in (53) with the system $(\overset{0}{x}^i)$ we get from (54)

$$\overset{0}{\alpha}^i_k = \frac{\partial x^i}{\partial \overset{0}{x}_k} = \delta^i_k - \Gamma^i_{rs}(P)\,.\,\overset{0}{\alpha}^r_k(x^s - x^s_P),$$

$$\frac{\partial \overset{0}{\alpha}^i_k}{\partial \overset{0}{x}^l} = -\Gamma^i_{rs}(P)\left[\frac{\partial \overset{0}{\alpha}^r_k}{\partial \overset{0}{x}^l}(x^s - x^s_P) + \overset{0}{\alpha}^r_k\,\overset{0}{\alpha}^s_l\right].$$

At the point P we thus have

$$\overset{0}{\alpha}^i_k(P) = \delta^i_k, \qquad \frac{\partial \overset{0}{\alpha}^i_k}{\partial \overset{0}{x}^l}(P) = -\Gamma^i_{rs}(P)\overset{0}{\alpha}^r_k\,\overset{0}{\alpha}^s_l, \tag{55}$$

and if we use (55) in (53) we get

$$\overset{0}{\Gamma}{}^i_{kl}(P) = 0 \tag{56}$$

at the point P. From (50) and (51) we then find that also $\overset{0}{\Gamma}_{i,kl}$ and $\partial \overset{0}{g}_{ik}/\partial \overset{0}{x}^l$ are zero at this point. Further, the components of the metric tensor g_{ik} at P are, on account of (55),

$$\overset{0}{g}_{ik}(P) = \overset{0}{\alpha}^l_i\,\overset{0}{\alpha}^m_k g_{lm}(P) = g_{ik}(P) \tag{57}$$

and the derivatives of the gravitational potentials $\overset{0}{\gamma}_i$ and $\overset{0}{\chi}$ defined by (VIII. 63, 94) are zero. Thus the gravitational acceleration is zero at P, the gravitational field has been locally transformed away. The system of reference $\overset{0}{R}$ corresponding to the coordinates $\overset{0}{x}^i$ is a *local system of inertia.*

The motion of the origin $\overset{0}{O}$ of the system $\overset{0}{R}$ relative to the original system R with coordinates (x^i) is obtained from (54) by putting $\overset{0}{x}^i = 0$, which gives

$$0 = x^i - x^i_P + \tfrac{1}{2}\Gamma^i_{rs}(P)(x^r - x^r_P)(x^s - x^s_P).$$

By differentiation of this equation with respect to an arbitrary parameter λ we get

$$0 = \frac{dx^\iota}{d\lambda} + \Gamma^\iota_{rs}(P)\frac{dx^r}{d\lambda}(x^s - x^s_P),$$

$$0 = \frac{d^2x^\iota}{d\lambda^2} + \Gamma^\iota_{rs}(P)\frac{d^2x^r}{d\lambda^2}(x^s - x^s_P) + \Gamma^\iota_{rs}(P)\frac{dx^r}{d\lambda}\frac{dx^s}{d\lambda}.$$

At the time $t = t_P = x^4_P/c$, corresponding to the event P, the space-time coordinates of $\overset{0}{O}$ are

$$x^\iota = x^\iota_P \quad \text{and} \quad \overset{0}{x}{}^\iota = 0.$$

At this moment the motion of $\overset{0}{O}$ is thus given by the equations

$$\frac{dx^\iota}{d\lambda} = 0, \qquad \frac{d^2x^\iota}{d\lambda^2} + \Gamma^\iota_{rs}(P)\frac{dx^r}{d\lambda}\frac{dx^s}{d\lambda} = 0, \tag{58}$$

which are the equations of motion of a freely falling particle which is momentarily at rest relative to R (cf. (49)).

Any geodesic line which goes through the point P, including the time track of freely falling particles and light rays, is in the system $\overset{0}{x}{}^\iota$ described by the equations (49) with the Christoffel symbol given by (56). Hence we get simply

$$\frac{d^2\overset{0}{x}{}^\iota}{d\lambda^2} = 0. \tag{59}$$

In a small region around P the local system of inertia thus has the same properties as a usual system of inertia. Without any change in the system of reference we may now further introduce local pseudo-Cartesian coordinates $\overset{0}{X}{}^\iota$ in $\overset{0}{R}$, by a transformation of the type (VIII. 59), so that the metric tensor $\overset{0}{G}_{\iota k}$ in this new system is

$$\overset{0}{G}_{\iota k}(P) = \begin{cases} 1 \text{ for } i = k = 1, 2, 3 \\ -1 \text{ for } i = k = 4 \\ 0 \text{ for } i \neq k \end{cases} \tag{60}$$

at the point P. The necessary transformation is even linear, therefore the derivatives of the metric tensor will remain zero at P, i.e.

$$\frac{\partial \overset{0}{G}_{\iota k}}{\partial \overset{0}{x}{}^l}(P) = 0. \tag{61}$$

In the following we shall always use such space-time coordinates in the local systems of inertia. By Lorentz transformations of the variables $\overset{0}{X}{}^\iota$ we can pass over to new local systems of inertia which in general will be moving relative to the original one.

In a small region around P, where terms of second order in $\overset{0}{X}{}^i$ can be neglected, the metric tensor may be regarded as constant in these systems. In accordance with the principle of equivalence, it is now assumed that *all laws of nature at the point P have the same form as in the special theory of relativity when expressed in terms of the local pseudo-Cartesian coordinates $\overset{0}{X}{}^i$.* By a simple transformation of coordinates we may then obtain these laws in a general covariant form. This requires the development of tensor analysis in a general Riemannian space.

105. Parallel displacement of vectors

Let a^i be the contravariant components of a vector at the point (x^i). By means of the Christoffel formulae it is now easily seen that the quantities $a^{*i} = a^i + d_p a^i$ with

$$d_p a^i = -\Gamma^i_{kl}\, dx^k a^l \tag{62}$$

transform like the contravariant components of a vector at the neighbouring point $(x^i + dx^i)$. From (53), (10), and (11) we obtain

$$\begin{aligned} d_p a'^i &= -\Gamma'^i_{kl}\, dx'^k a'^l = -\Gamma'^i_{kl}\, \alpha^k_m \alpha^l_n\, dx^m a^n \\ &= \left(-\alpha^i_r \frac{\partial \breve{\alpha}^r_k}{\partial x'^l} \alpha^l_n \alpha^k_m - \alpha^i_r\, \Gamma^r_{mn}\right) dx^m a^n. \end{aligned} \tag{63}$$

Further, we get by differentiation of (11)

$$\frac{\partial \breve{\alpha}^i_l}{\partial x^n} \alpha^l_k + \breve{\alpha}^i_l \frac{\partial \alpha^l_k}{\partial x^n} = 0 \tag{11'}$$

and the first term in the brackets may therefore be written

$$-\alpha^i_r \frac{\partial \breve{\alpha}^r_k}{\partial x'^l} \frac{\partial x'^l}{\partial x^n} \alpha^k_m = -\alpha^i_r \frac{\partial \breve{\alpha}^r_k}{\partial x^n} \alpha^k_m = \alpha^i_r \breve{\alpha}^r_k \frac{\partial \alpha^k_m}{\partial x^n} = \frac{\partial \alpha^i_m}{\partial x^n} = \frac{\partial \alpha^i_n}{\partial x^m},$$

where we have made use of (11) and (14).

Hence (63) becomes

$$d_p a'^i = \frac{\partial \alpha^i_n}{\partial x^m}\, dx^m a^n + \alpha^i_r\, d_p a^r,$$

and if we neglect terms of second order in dx^i the transformation law for the quantities $a^i + d_p a^i$ can be written

$$\begin{aligned} a'^i + d_p a'^i &= \left(\alpha^i_n(x) + \frac{\partial \alpha^i_n}{\partial x^m}\, dx^m\right) a^n + \alpha^i_n\, d_p a^n \\ &= \alpha^i_n(x+dx)(a^n + d_p a^n), \end{aligned} \tag{64}$$

where $\alpha^i_n(x+dx)$ are the transformation coefficients taken at the point $(x^i + dx^i)$. (64) shows that the quantities $a^{*i} = a^i + d_p a^i$ are the contravariant components of a vector at the point $(x^i + dx^i)$. In a flat space, the vector a^{*i} is identical with the vector obtained by parallel displacement of the vector a^i from the point (x^i) to the neighbouring point (x^i+dx^i); for if we introduce pseudo-Cartesian coordinates, the Christoffel symbols and therefore also $d_p a^i$ vanish and the components a^{*i} and a^i of the two vectors are equal in this system. If we use a curvilinear system of coordinates in a flat space, the contravariant components of the two vectors will, thus, differ by the amount $d_p a^i$ defined by (62).

It is now natural also in a general Riemannian space to *define* the parallel displacement of a vector by the equation (62).

If $\overset{0}{X}{}^i$ is a local system of inertia for the point considered, the components a^i of a vector at this point are unchanged by parallel displacements exactly as in a pseudo-Cartesian system of coordinates in a pseudo-Euclidean space.

From (62) it follows that the norm of a vector, and also the scalar product of two vectors a^i and b^i at the point (x^i), are unchanged by a parallel displacement, for from (62) we get

$$d_p(g_{ik}a^ib^k) = \frac{\partial g_{ik}}{\partial x^l}\,dx^la^ib^k - g_{ik}\,\Gamma^i_{ls}\,dx^la^sb^k - g_{ik}\,a^i\Gamma^k_{ls}\,dx^lb^s$$

$$= dx^la^ib^k\left\{\frac{\partial g_{ik}}{\partial x^l} - g_{rk}\,\Gamma^r_{il} - g_{ir}\,\Gamma^r_{kl}\right\} = 0 \tag{65}$$

on account of the identity

$$\frac{\partial g_{ik}}{\partial x^l} - g_{rk}\,\Gamma^r_{il} - g_{ir}\,\Gamma^r_{kl} \equiv 0 \tag{66}$$

following from (50) and (51).

Using one of the other forms (24) in which the scalar product can be written, we get

$$0 = d_p(a_i\,b^i) = d_p\,a_i\,.\,b^i - a_i\,\Gamma^i_{kl}\,dx^kb^l$$

$$\equiv (d_p\,a_i - \Gamma^l_{ik}\,dx^ka_l)b^i.$$

Since this equation must hold for an arbitrary vector b^i, we get for the change in the covariant components of a vector by a parallel displacement

$$d_p\,a_i \doteq \Gamma^l_{ik}\,dx^ka_l = \Gamma_{l,ik}\,dx^ka^l. \tag{67}$$

Finally, expressing the scalar product in the form $g^{ik}a_i\,b_k$, we get by means of (67)

$$0 = d_p(g^{ik}a_i\,b_k) = dx^la_i\,b_k\left(\frac{\partial g^{ik}}{\partial x^l} + g^{ir}\Gamma^k_{lr} + g^{kr}\Gamma^i_{lr}\right)$$

and, since this equation must hold for arbitrary dx^i, a_ι, b_ι, the equation

$$\frac{\partial g^{ik}}{\partial x^l}+g^{ir}\Gamma^k_{lr}+g^{kr}\Gamma^i_{lr} \equiv 0 \tag{68}$$

must be identically satisfied.

Multiplying (66) by g^{ik} and applying (7), we obtain

$$\Gamma^k_{ik} = \tfrac{1}{2}g^{kl}\frac{\partial g_{kl}}{\partial x^i} = -\tfrac{1}{2}g_{kl}\frac{\partial g^{kl}}{\partial x^i}. \tag{68'}$$

Further we get by means of a well-known theorem the derivative of the determinant $g = |g_{ik}|$ in the form

$$\frac{\partial g}{\partial x^i} = A_{kl}\frac{\partial g_{kl}}{\partial x^i} = gg^{kl}\frac{\partial g_{kl}}{\partial x^i} = -gg_{kl}\frac{\partial g^{kl}}{\partial x^i}, \tag{69'}$$

where we have made use of (6) and (7).

Hence
$$\Gamma^k_{ik} = \frac{1}{2g}\frac{\partial g}{\partial x^i} = \frac{\partial}{\partial x^i}\ln\sqrt{|g|}, \tag{69}$$

a relation which we shall use later.

Consider again a geodesic line defined by the equation (49). If λ is an invariant parameter,

$$U^i = \frac{dx^i}{d\lambda} \tag{70}$$

is a four-vector lying in the direction of the tangent. The equation (49) may then be written

$$\frac{dU^i}{d\lambda} = -\Gamma^i_{kl}\frac{dx^k}{d\lambda}U^l. \tag{71}$$

A comparison of (71) with (62) shows that the different vectors U^i along the geodesic line are obtained by parallel displacement along this line, a property which they have in common with the straight lines in a Euclidean space. The geodesic line connecting two points is thus not only the line with a stationary value of length, it is also the 'straightest' line. The norm of the vector U^i must therefore be constant along the line, i.e. $g_{ik}U^iU^k$ is independent of λ, in accordance with (VIII. 31).

If this property of the geodesic lines is expressed in terms of the covariant components U_i of the vector $dx^i/d\lambda$, we get, according to (67) and (51),

$$\frac{dU_i}{d\lambda} = \Gamma_{l,ik}U^kU^l = \tfrac{1}{2}(\Gamma_{l,ik}+\Gamma_{k,il})U^kU^l = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}U^kU^l. \tag{72}$$

This equation is identical with the equation (48).

The equations (62) define the change in the components a^i of a vector by an infinitesimal parallel displacement along the vector (dx^i). The total change of a^i by a displacement along a finite curve may thus be obtained by integration. In a flat space the total change of a^i by parallel displacement along a closed curve will be zero. This is seen at once if we use a Cartesian or pseudo-Cartesian system of coordinates, for in this system the components a^i are not changed at all by the displacement. The final vector a^{*i} obtained by the displacement along the closed curve is thus equal to the original vector a^i, and this must then also be true if we afterwards introduce curvilinear coordinates. In a curved space, however, the final vector a^{*i} will in general be different from a^i, the difference $a^{*i}-a^i$ depending on the closed curve (see § 107). Thus, if a vector is given a parallel displacement from a point P_1 to a point P_2 along a certain curve connecting P_1 and P_2, the resulting vector a^{*i} will depend on the form of this curve if the space is curved, while it is independent of the curve in a flat space. This is in fact the only essential difference between curved and flat space.

106. Tensor analysis. Covariant differentiation

As in the case of the flat pseudo-Euclidean space of the special theory of relativity, we speak of a tensor (or pseudo-tensor) field of rank n in a general Riemannian space if a tensor (or pseudo-tensor) of rank n is connected with every point in this space. Similarly as in § 48, we can derive from such a field a new tensor field of rank $n+1$ by a differentiation process. From a tensor field of rank 0, i.e. from a scalar field

$$\phi'(x') = \phi(x), \tag{73}$$

we can thus derive a vector field $\operatorname{grad}\phi$ with the covariant components

$$\operatorname{grad}_i\phi = \frac{\partial\phi}{\partial x^i}. \tag{74}$$

By means of (73) and (10) we get at once

$$\frac{\partial\phi'}{\partial x'^i} = \frac{\partial\phi}{\partial x^k}\frac{\partial x^k}{\partial x'^i} = \bar{\alpha}_i^k\frac{\partial\phi}{\partial x^k}, \tag{75}$$

which shows that the $\partial\phi/\partial x^i$ are in fact the covariant components of a vector.

However, if we try in the same way to form a tensor field of rank 2 from a vector field a^i by differentiation of the transformation equations (15), we get

$$\frac{\partial a'^i}{\partial x'^k} = \frac{\partial\alpha_l^i}{\partial x'^k}a^l + \alpha_l^i\,\bar{\alpha}_k^m\frac{\partial a^l}{\partial x^m}. \tag{76}$$

From this equation we see that the $\partial a^i/\partial x^k$ will be the mixed components of a tensor only if the coefficients α_l^i are constant. On the other hand, we get by means of the Christoffel formulae (53) and the relations (11), (14), and (11′) (p. 276)

$$\Gamma'^{\iota}_{kr}\,a'^{r} = \left(-\frac{\partial\alpha^{\iota}_{l}}{\partial x'^{k}}+\alpha^{\iota}_{r}\,\check{\alpha}^{s}_{k}\,\Gamma^{r}_{sl}\right)a^{l}. \tag{77}$$

By addition of (76) and (77) we find that the quantities

$$a^{\iota}{}_{,k} = \frac{\partial a^{\iota}}{\partial x^{k}}+\Gamma^{\iota}_{kr}\,a^{r} \tag{78}$$

transform according to the law

$$a'^{\iota}{}_{,k} = \frac{\partial a'^{\iota}}{\partial x'^{k}}+\Gamma'^{\iota}_{kr}\,a'^{r} = \alpha^{\iota}_{l}\,\check{\alpha}^{m}_{k}\,a^{l}_{,m}.$$

Thus, by the process of 'covariant differentiation' (78) of the contravariant components a^{ι} of a vector we get the mixed components of a tensor field of rank 2.

This process may be described geometrically in the following way. Let $a^{\iota}(P)$ and $a^{\iota}(P')$ be the vectors of the vector field connected with the neighbouring points P and P' with coordinates (x^{ι}) and $(x^{\iota}+dx^{\iota})$, respectively. The differences

$$da^{\iota} = a^{\iota}(P')-a^{\iota}(P) = \frac{\partial a^{\iota}}{\partial x^{k}}\,dx^{k}$$

will then not be the components of a vector, since $a^{\iota}(P)$ and $a^{\iota}(P')$ belong to different points. However, if $a^{*\iota}(P')$ denotes the vector obtained by parallel displacement of a^{ι} from P to P', the difference

$$\begin{aligned} a^{\iota}(P')-a^{*\iota}(P') &= a^{\iota}(P')-a^{\iota}(P)-(a^{*\iota}(P')-a^{\iota}(P)) \\ &= \left(\frac{\partial a^{\iota}}{\partial x^{k}}+\Gamma^{\iota}_{kr}\,a^{r}\right)dx^{k} = a^{\iota}{}_{,k}\,dx^{k} \end{aligned}$$

will represent an infinitesimal vector at the point P'. Neglecting terms of the second order in (dx^{k}), $a^{\iota}{}_{,k}\,dx^{k}$ will also be a vector at P, and since this must hold for arbitrary infinitesimal vectors dx^{ι}, the quantities $a^{\iota}{}_{,k}$ must be the mixed components of a tensor of rank 2.

If this consideration is applied to the covariant components a_{ι} of the vector field, we find, using the formulae (67) instead of (62) for the parallel displacement, that the quantities

$$a_{\iota,k} \equiv \frac{\partial a_{i}}{\partial x^{k}}-\Gamma^{r}_{\iota k}\,a_{r} \tag{79}$$

are the covariant components of a tensor. In a geodesic system and, in particular, in a Cartesian system of coordinates in a flat space, where the Christoffel symbols are zero, the covariant differentiations (78) and (79) are ordinary differentiations.

Let us now consider the special case, where we have only a vector $a^i(\lambda)$ connected with each point on a curve with the parametric representation

$$x^i = x^i(\lambda).$$

We can then define the *covariant derivative* of this vector with respect to λ by

$$\left.\begin{aligned} \frac{Da^i}{d\lambda} &= \frac{da^i}{d\lambda} + \Gamma^i_{kl} U^k a^l \\ U^i &= \frac{dx^i}{d\lambda} \end{aligned}\right\}. \tag{80}$$

The quantity $Da^i(\lambda)/d\lambda$ is obviously a vector at the point P with the coordinates $x^i(\lambda)$; for from (62) we see that

$$\frac{Da^i}{d\lambda} = \lim_{P'\to P} \frac{[a^i(P')-a^i(P)-(a^{*i}(P')-a^i(P))]}{\Delta\lambda} = \lim_{P'\to P} \frac{a^i(P')-a^{*i}(P')}{\Delta\lambda}, \tag{80'}$$

where P' with the coordinates $x^i(\lambda+\Delta\lambda)$ in this limiting process is approaching the point P along the curve. Similarly we get for the *covariant derivative* of the covariant components a_i

$$\frac{Da_i}{d\lambda} = \frac{da_i}{d\lambda} - \Gamma_{l,ik} U^k a^l = \frac{da_i}{d\lambda} - \Gamma^l_{ik} U^k a_l. \tag{80''}$$

By means of (51) we get

$$\begin{aligned} \frac{Da_i}{d\lambda} &= \frac{d(g_{ik}a^k)}{d\lambda} - \Gamma_{l,ik} U^k a^l = \frac{dg_{ik}}{d\lambda} a^k + g_{ik}\frac{da^k}{d\lambda} - \Gamma_{l,ik} U^k a^l \\ &= \left(\frac{\partial g_{ik}}{\partial x^l} - \Gamma_{k,il}\right) U^l a^k + g_{ik}\frac{da^k}{d\lambda} = g_{ik}\frac{da^k}{d\lambda} + \Gamma_{i,kl} U^l a^k \\ &= g_{ik}\left(\frac{da^k}{d\lambda} + \Gamma^k_{lm} U^l a^m\right) = g_{ik}\frac{Da^k}{d\lambda}, \end{aligned} \tag{80'''}$$

i.e. $Da_i/d\lambda$ and $Da^i/d\lambda$ are the covariant and contravariant components of the *same* vector.

According to (71) and (72) the equations for the time track of a free particle can now also be written

$$\frac{DU^i}{d\lambda} = 0 \quad \text{or} \quad \frac{DU_i}{d\lambda} = 0. \tag{80^{iv}}$$

The process of covariant differentiation may be applied also to tensor fields of higher rank. Consider, for instance, a tensor field of rank 2 with contravariant components t^{ik}. Since each index is transformed separately according to the same rule as for a vector, it is clear that the quantities

$$t^{ik}{}_{,l} = \frac{\partial t^{ik}}{\partial x^l} + \Gamma^i_{lr} t^{rk} + \Gamma^k_{lr} t^{ir}, \tag{81}$$

with two contravariant indices and one covariant index, are the mixed components of a tensor of rank 3. This may be verified in the same way as for a vector field by means of the transformation laws of tensors and Christoffel symbols. Similarly it is seen that

$$\left.\begin{aligned} t_{ik,l} &= \frac{\partial t_{ik}}{\partial x^l} - \Gamma^r_{il} t_{rk} - \Gamma^r_{kl} t_{ir} \\ t^i{}_{k,l} &= \frac{\partial t^i{}_k}{\partial x^l} + \Gamma^i_{lr} t^r{}_k - \Gamma^r_{kl} t^i{}_r \end{aligned}\right\} \tag{82}$$

are purely covariant and mixed components, respectively, of a tensor of rank 3.

These rules may be extended to the case of a tensor field of rank n, the number of terms containing Christoffel symbols being n in this case, and if we put

$$\phi_{,i} = \frac{\partial \phi}{\partial x^i} \tag{83}$$

in the case of a scalar field, this rule also applies to tensor fields of rank 0. Since the scalar product $a_i b^i$ of two vector fields a^i, b^i is a scalar field, we have

$$\begin{aligned} (a_i b^i)_{,k} &= \frac{\partial a_i}{\partial x^k} b^i + a_i \frac{\partial b^i}{\partial x^k} = \left(\frac{\partial a_i}{\partial x^k} - \Gamma^r_{ik} a_r\right) b^i + a_i \left(\frac{\partial b^i}{\partial x^k} + \Gamma^i_{kr} b^r\right) \\ &= a_{i,k} b^i + a_i b^i{}_{,k}, \end{aligned} \tag{84}$$

i.e. the usual rule for differentiation of a product holds also for covariant differentiation. This rule is easily seen to hold for the contracted product of any two tensors of arbitrary rank; for instance we have

$$\left.\begin{aligned} (t_{ik} a^k)_{,l} &= t_{ik,l} a^k + t_{ik} a^k{}_{,l} \\ (t^{ik} a_k)_{,l} &= t^{ik}{}_{,l} a_k + t^{ik} a_{k,l} \end{aligned}\right\}. \tag{85}$$

The identities (66) and (68) may now be written

$$g_{ik,l} \equiv 0, \qquad g^{ik}{}_{,l} \equiv 0, \tag{86}$$

i.e. the covariant derivatives of the metric tensor are zero. By covariant differentiation of (16) and (17) we therefore get

$$a_{i,l} = g_{ik,l} a^k + g_{ik} a^k{}_{,l} = g_{ik} a^k{}_{,l}, \qquad a^i{}_{,l} = g^{ik} a_{k,l}, \tag{87}$$

i.e. the quantities $a_{i,k}$ and $a^i{}_{,k}$ are components of the *same* tensor of rank 2. In the same way we see that the quantities $t^{ik}{}_{,l}$, $t^i{}_{k,l}$, and $t_{ik,l}$, defined by (81) and (82), are components of the *same* tensor of rank 3.

To obtain a generalization of the differential operators defined in § 48 for the special case of pseudo-Euclidean space, we simply have to replace the differentiations in § 48 by covariant differentiations. For the curl of a vector field a_i we thus get

$$\mathrm{curl}_{ik}\{a_i\} = a_{k,i} - a_{i;k} = \frac{\partial a_k}{\partial x^i} - \frac{\partial a_i}{\partial x^k}, \tag{88}$$

since the terms containing the Christoffel symbols cancel.

The covariant expression for the divergence of a vector is obtained by contraction of the tensor $a^i{}_{,k}$, i.e.

$$\mathrm{div}\{a^i\} = a^i{}_{,i} = \frac{\partial a^i}{\partial x^i} + \Gamma^i_{ir} a^r. \tag{89}$$

On account of (69) this may also be written

$$\mathrm{div}\{a^i\} = \frac{\partial a^i}{\partial x^i} + \frac{a^i}{\sqrt{|g|}} \frac{\partial}{\partial x^i}\sqrt{|g|} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^i}(\sqrt{(|g|)}a^i). \tag{90}$$

In the $(3+1)$-space g is negative, i.e. $|g| = -g$, for a space with positive definite metric tensor $g > 0$ and $|g| = g$. The covariant expression for d'Alembert's operator is now

$$\Box\psi = \mathrm{div}\{\mathrm{grad}\,\psi\} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^i}\left(\sqrt{(|g|)}g^{ik}\frac{\partial\psi}{\partial x^k}\right). \tag{91}$$

The contravariant components of the divergence of a tensor field T^{ik} are

$$\mathrm{div}^i\{T^{ik}\} = T^{ik}{}_{,k} = \frac{\partial T^{ik}}{\partial x^k} + \Gamma^i_{kr} T^{rk} + \Gamma^k_{kr} T^{ir}$$

$$= \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^k}(\sqrt{(|g|)}T^{ik}) + \Gamma^i_{kl} T^{kl}, \tag{92 a}$$

and similarly we get for the covariant components

$$\mathrm{div}_i\{T_i{}^k\} = T_i{}^k{}_{,k} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^k}(\sqrt{(|g|)}T_i^{\,k}) - \Gamma_{r,is} T^{rs}. \tag{92 b}$$

For a symmetrical tensor this reduces to

$$\mathrm{div}\{T^k_i\} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^k}(\sqrt{(|g|)}T^k_i) - \frac{1}{2}\frac{\partial g_{rs}}{\partial x^i} T^{rs} \tag{92 c}$$

on account of (51).

For an antisymmetrical tensor F^{ik}, the last term in (92 a) is zero on account of the symmetry of Γ^i_{kl} in the lower indices; hence

$$\mathrm{div}^i\{F^{ik}\} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^k}(\sqrt{(|g|)}F^{ik}). \tag{93}$$

Further, we get for the curl of an antisymmetrical tensor

$$\mathrm{curl}_{ikl}\{F_{ik}\} = F_{ik,l}+F_{kl,i}+F_{li,k} = \frac{\partial F_{ik}}{\partial x^l}+\frac{\partial F_{kl}}{\partial x^i}+\frac{\partial F_{li}}{\partial x^k} \tag{94}$$

since the terms containing the Christoffel symbols cancel.

The generalization of Gauss's theorem (IV. 191) is obvious. Using (43), (47), and (90), we get for an arbitrary vector field a^i in the (3+1)-space

$$\int\limits_{\Sigma} \mathrm{div}\{a\}\, d\Sigma = \int \frac{\partial}{\partial x^i}(\sqrt{(-g)}a^i)\, dx^1dx^2dx^3dx^4$$

$$= \int\limits_{V} a^i\, dV_i = \int a^i\sqrt{(-g)}\, \delta_{iklm}\, dx^k\delta x^l\Delta x^m, \tag{95}$$

where (dx^k), (δx^l), (Δx^m) are three infinitesimal vectors in the three-dimensional boundary V of the four-dimensional region.

107. The curvature tensor

Let a_k be an arbitrary vector field and $a_{k,l}$ the tensor of rank 2 obtained by covariant differentiation. It may be written in the two forms

$$a_{k,l} = \frac{\partial a_k}{\partial x^l} - \Gamma^r_{kl}a_r = \frac{\partial a_k}{\partial x^l} - \Gamma_{r,kl}a^r. \tag{96}$$

Similarly the tensor of rank 3 obtained by a further covariant differentiation can be written in the two forms

$$a_{k,lm} = \frac{\partial a_{k,l}}{\partial x^m} - \Gamma^r_{km}\, a_{r,l} - \Gamma^r_{lm}\, a_{k,r} = \frac{\partial a_{k,l}}{\partial x^m} - \Gamma_{r,km}\, a^r{}_{,l} - \Gamma^r_{lm}\, a_{k,r}, \tag{97}$$

where $a^r{}_{;l}$ is obtained from (78) or (87).

Using the first expression (96) in the first form of $a_{k,lm}$ in (97), we get an expression for $a_{k,lm}$ which is a linear function of a_i and its first and second derivatives. However, if we subtract the tensor $a_{k,ml}$ obtained by interchanging the order of the covariant differentiations, the derivatives of a_i disappear and we get simply

$$a_{k,lm} - a_{k,ml} = -R^i{}_{klm}\, a_i, \tag{98}$$

where the coefficients of a_i are given by

$$R^i{}_{klm} = \frac{\partial \Gamma^i_{kl}}{\partial x^m} - \frac{\partial \Gamma^i_{km}}{\partial x^l} + \Gamma^i_{rm}\, \Gamma^r_{kl} - \Gamma^i_{rl}\, \Gamma^r_{km}. \tag{99}$$

Using instead the second expressions (96) and (97), we get in the same way

$$a_{k;lm} - a_{k,ml} = -R_{iklm}\, a^i, \tag{100}$$

where

$$R_{iklm} = \frac{\partial \Gamma_{i,kl}}{\partial x^m} - \frac{\partial \Gamma_{i,km}}{\partial x^l} + \Gamma^r_{il}\,\Gamma_{r,km} - \Gamma^r_{im}\,\Gamma_{r,kl}$$

$$= \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial x^k \partial x^m} + \frac{\partial^2 g_{km}}{\partial x^i \partial x^l} - \frac{\partial^2 g_{im}}{\partial x^k \partial x^l} - \frac{\partial^2 g_{kl}}{\partial x^i \partial x^m}\right) +$$

$$+ g^{rs}(\Gamma_{r,il}\,\Gamma_{s,km} - \Gamma_{r,im}\,\Gamma_{s,kl}). \quad (101)$$

Since the left-hand sides of (98) and (100) are the components of a tensor for any vector field a_i, the quantities $R^i{}_{klm}$ and R_{iklm} must be the components of the same tensor of rank 4, i.e.

$$R^i{}_{klm} = g^{in} R_{nklm}. \quad (102)$$

The tensor R_{iklm} is called the *Riemann–Christoffel curvature tensor*. The equations (98) represent the commutation law for the covariant differentiation of a vector field. The corresponding law for the covariant differentiation of a tensor field t_{ik} is easily seen to be

$$t_{kl,mn} - t_{kl,nm} = -R^i{}_{kmn}\, t_{il} - R^i{}_{lmn}\, t_{ki}. \quad (103)$$

The geometrical meaning of the curvature tensor becomes apparent if one considers a parallel displacement of a vector a^i along the contour of an infinitesimal parallelogram defined by two infinitesimal vectors (dx^i), (δx^i). As mentioned in § 105, the vector a^{*i} resulting from this process will in general differ from the vector a^i. By means of the parallel displacement laws (62) and (67) it may now be verified by elementary calculations that the differences

$$\Delta a^i = a^{*i} - a^i, \qquad \Delta a_i = a^*_i - a_i$$

between the components of these vectors are given by

$$\Delta a^i = \tfrac{1}{2} R^i{}_{klm}\, a^k\, d\sigma^{lm}, \qquad \Delta a_i = \tfrac{1}{2} R_{iklm}\, a^k\, d\sigma^{lm}, \quad (104)$$

where

$$d\sigma^{lm} = dx^l \delta x^m - dx^m \delta x^l.$$

In a flat space, where we can introduce a system of coordinates in which the components of the metric tensor are constant, we have obviously

$$R_{iklm} = 0. \quad (105)$$

This equation is thus a necessary condition for the space to be a flat space. It is, however, also a sufficient condition; for if (105) holds at all points, it is possible to find a transformation (9) which makes the transformed components g'_{ik} independent of the space-time variables (x'^i) (see Appendix 6).

From (101) it follows immediately that the curvature tensor satisfies the relations

$$R_{iklm} = -R_{ikml} = -R_{kilm} = R_{lmik}, \tag{106a}$$

$$R_{iklm}+R_{ilmk}+R_{imkl} = 0. \tag{106b}$$

Besides these algebraic relations the curvature tensor satisfies a set of differential identities which may be obtained in the following way.† On account of the general rule for covariant differentiation of a contracted product, we get from (98) by covariant differentiation

$$a_{k,lmn}-a_{k;mln} = -R^{i}{}_{klm;n}\,a_{i}-R^{i}{}_{klm}\,a_{i,n}.$$

Adding to this equation the two equations obtained by cyclic permutation of the indices l, m, n, we get

$$\begin{aligned}(a_{k,lmn}-a_{k,lnm})+(a_{k;mnl}-a_{k,mln})+(a_{k,nlm}-a_{k,nml}) \\ = -(R^{i}{}_{klm,n}+R^{i}{}_{kmn,l}+R^{i}{}_{knl,m})a_{i}- \\ -(R^{i}{}_{klm}\,a_{i,n}+R^{i}{}_{kmn}\,a_{i,l}+R^{i}{}_{knl}\,a_{i,m}).\end{aligned} \tag{107}$$

Each of the three brackets on the left-hand side may be transformed by means of the equations (103) applied to the tensors $a_{k,l}$, $a_{k;m}$, and $a_{k,n}$, respectively. In this way we get six terms, three of which cancel on account of (106b), while the other three cancel with the terms in the second bracket on the right-hand side of (107). Hence we get

$$(R^{i}{}_{klm,n}+R^{i}{}_{kmn,l}+R^{i}{}_{knl,m})a_{i} = 0.$$

Since this equation must hold for an arbitrary vector field a_{i} we are led to the *Bianchi identities*

$$R^{i}{}_{klm,n}+R^{i}{}_{kmn,l}+R^{i}{}_{knl,m} \equiv 0. \tag{108}$$

On account of the identities (106) the number of algebraically independent components of the curvature tensor is 20 in a four-dimensional space, 6 in a three-dimensional space, and 1 in a two-dimensional space.

108. The contracted forms of the curvature tensor

By contraction of the tensor $R^{i}{}_{klm}$ of rank 4 we get a tensor of rank 2 which, on account of (106a), may be written in the different forms

$$R_{ik} = R^{r}{}_{irk} = -R^{r}{}_{ikr} = -R_{i}{}^{r}{}_{rk} = R_{rk}{}^{r}{}_{i} = R^{r}{}_{kri}. \tag{109}$$

As the last expression is equal to the first with i and k interchanged, this *contracted curvature tensor* is obviously symmetrical

$$R_{ik} = R_{ki}. \tag{110}$$

By further contraction we get the curvature scalar

$$R = R^{i}_{i} = g^{ik}R_{ik}. \tag{111}$$

† See, e.g., P. G. Bergmann, *Introduction to the Theory of Relativity*, New York, 1942, p. 169.

Contracting the Bianchi identities (108) with respect to the indices i and l, and using the relations (109) and (106a), we get

$$R_{km,n}+R^{i}{}_{kmn,i}-R_{kn;m}=0 \quad \text{or} \quad R^{k}_{m,n}-R^{ki}{}_{mn;i}-R^{k}_{n,m}=0.$$

Further contraction with respect to the indices k and m gives

$$R_{,n}-2R^{k}_{n,k}=0$$

or, by multiplication by g^{in} and application of (86) and of the general rule for covariant differentiation of products,

$$(R^{ik}-\tfrac{1}{2}g^{ik}R)_{,k}=0. \tag{112}$$

This equation expresses the fact that the covariant divergence of the symmetrical tensor

$$R^{ik}-\tfrac{1}{2}g^{ik}R \tag{113}$$

is zero. On account of the symmetry property, this tensor has ten independent components only.

From (109), (99), and (69) we get the following explicit expressions for R_{ik}:

$$\begin{aligned} R_{ik} &= \frac{\partial \Gamma^{l}_{il}}{\partial x^{k}}-\frac{\partial \Gamma^{l}_{ik}}{\partial x^{l}}+\Gamma^{r}_{il}\Gamma^{l}_{kr}-\Gamma^{r}_{ik}\Gamma^{l}_{lr} \\ &= \frac{\partial^{2}\ln\sqrt{|g|}}{\partial x^{i}\partial x^{k}}-\frac{\partial \Gamma^{l}_{ik}}{\partial x^{l}}+\Gamma^{r}_{il}\Gamma^{l}_{kr}-\Gamma^{r}_{ik}\frac{\partial \ln\sqrt{|g|}}{\partial x^{r}}. \end{aligned} \tag{114}$$

X

THE INFLUENCE OF GRAVITATIONAL FIELDS ON PHYSICAL PHENOMENA

109. Mechanics of free particles in the presence of gravitational fields

By means of the formalism of the general tensor calculus and the assumption made at the end of § 104, based on the principle of equivalence, the physical laws of the special theory of relativity can now be generalized in an unambiguous way. Since the tensor equations of the special theory are assumed to hold in a local system of inertia, i.e. in a geodesic system of space-time coordinates, the problem of finding, for instance, the fundamental equations of mechanics and electrodynamics in the presence of gravitational fields reduces to a purely geometrical problem in 4-space.

Let us first consider the motion of a particle in an arbitrary system of coordinates (x^i). Let

$$x^i = x^i(\tau) \tag{1}$$

be the equation of the time track of the motion, τ being the proper time of the particle measured by a standard clock following the particle. The contravariant components of the four-velocity are then, on account of (VIII. 98),

$$U^i \equiv \frac{dx^i}{d\tau} = (\Gamma u^\iota, c\Gamma), \tag{2}$$

where $u^\iota = dx^\iota/dt$ are the contravariant components of the spatial velocity and

$$\frac{1}{\Gamma} = \left[\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}^2 - \frac{u^2}{c^2}\right]^{\frac{1}{2}} \tag{3}$$

is the analogue of the Lorentz factor in the presence of a gravitational field with the dynamical potentials (γ_ι, χ). In a local system of inertia, the expressions (2) become equivalent to (IV. 39).

The covariant components of the four-velocity vector are by (IX. 16) and (VIII. 63, 64, 94)

$$\left.\begin{aligned} U_i &= g_{ik} U^k = g_{\iota\kappa} u^\kappa \Gamma + g_{\iota 4} c\Gamma \\ \text{i.e.}\quad U_\iota &= \Gamma\left[u_\iota - \gamma_\iota(\gamma_\kappa u^\kappa) + c\gamma_\iota\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\right] \\ U_4 &= -\Gamma c\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\} \end{aligned}\right\}, \tag{4}$$

where $u_\iota = \gamma_{\iota\kappa} u^\kappa$ are the covariant components of the three-dimensional velocity vector, calculated from the contravariant components by means of the spatial metric tensor $\gamma_{\iota\kappa}$. By purely spatial transformations

$$x'^\iota = x'^\iota(x^\kappa), \qquad x'^4 = x^4 \tag{5}$$

the spatial parts U^ι and U_ι of the four-velocity transform like the contravariant and covariant components, respectively, of a vector; but unless the system of coordinates (x^ι) is time-orthogonal, U^ι and U_ι will represent different space vectors. According to (IV. 41) we have of course

$$U_\iota U^\iota = -c^2, \tag{6}$$

which also follows from the explicit expressions (2) and (4) for U^ι and U_ι.

The components of the four-acceleration in curvilinear coordinates are obviously

$$A^\iota = \frac{DU^i}{d\tau} = \frac{dU^\iota}{d\tau} + \Gamma^i_{kl} U^k U^l, \qquad A_\iota = \frac{DU_\iota}{d\tau} = \frac{dU_\iota}{d\tau} - \Gamma_{l,\iota k} U^k U^l \tag{7}$$

obtained by covariant differentiation (IX. 80, 80″) of U^ι and U_i with respect to the proper time. In a local system of inertia, (7) reduces to (IV. 42). According to (IV. 41′) we have

$$U_\iota A^\iota = U^\iota A_\iota = 0. \tag{7′}$$

The contravariant and covariant components of the four-momentum vector are now, by (IV. 50),

$$P^\iota = \mathring{m}_0 U^\iota, \qquad P_\iota = \mathring{m}_0 U_\iota, \tag{8}$$

where $\mathring{m}_0$ is the proper mass of the particle, i.e. the mass measured in a local rest system of inertia. $\mathring{m}_0$ is equal to the rest mass which the particle would have when placed in a usual system of inertia.

For a free particle, i.e. a particle which is acted upon by gravitational forces only, we have in a local system of inertia $(\mathring{x}^\iota)$

$$\frac{d\mathring{P}^\iota}{d\tau} = 0, \qquad \frac{d\mathring{P}_\iota}{d\tau} = 0.$$

In a general system (x^ι) these equations take the forms

$$\frac{DP^\iota}{d\tau} = 0, \qquad \frac{DP_\iota}{d\tau} = 0, \tag{9}$$

where

$$\frac{DP^\iota}{d\tau} = \frac{dP^\iota}{d\tau} + \Gamma^i_{kl} U^k P^l, \tag{10a}$$

$$\frac{DP_\iota}{d\tau} = \frac{dP_i}{d\tau} - \Gamma_{l,\iota k} U^k P^l \tag{10b}$$

represent the covariant derivatives of the four-momentum vector.

According to (IX. 80′) the quantities DP^i are equal to the differences of the components of the momentum four-vector $P^i(\tau+d\tau)$ connected with the point $x^i(\tau+d\tau)$ and the vector P^{*i} obtained by parallel displacement through the distance $dx^i = U^i\,d\tau$ of the vector $P^i(\tau)$ connected with the point $x^i(\tau)$. Thus the equations (9) express the fact that the four-momentum vector at the time $t+d\tau$ is obtained from the vector at the time τ by a parallel displacement. On account of (IX. 51), the last equations (9) may also be written

$$\frac{DP_i}{d\tau} = \frac{dP_i}{d\tau} - \frac{\mathring{m}_0}{2}(\Gamma_{l,ik}+\Gamma_{k,il})U^kU^l = \frac{dP_i}{d\tau} - \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}U^kP^l = 0. \tag{11}$$

110. Momentum and mass of a particle. Gravitational force

The equations (9) or (11) determine the motion of a material particle in a given external gravitational field. Strictly speaking, the particle itself will create a gravitational field which should also be described by the functions g_{ik}. In the present sections we assume, however, that this field is weak in comparison with the external field so that its influence on g_{ik} may be neglected. The g_{ik} may then be regarded as known functions of the space-time coordinates (x^i).

In order to provide a better understanding of the physical meaning of the quantities occurring in (11) we try to write these equations in the form of three-dimensional vector equations. Let us define a spatial vector p_ι, p^ι by the equations

$$\left.\begin{aligned} p^\iota &= mu^\iota \\ p_\iota &= mu_\iota = m\gamma_{\iota\kappa}u^\kappa = \gamma_{\iota\kappa}p^\kappa \\ m = \mathring{m}_0\,\Gamma &= \mathring{m}_0\left[\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}^2 - \frac{u^2}{c^2}\right]^{-\frac{1}{2}} \end{aligned}\right\}. \tag{12}$$

If $p_\iota = p_\iota(t)$ is regarded as a function of the time variable t we can construct a new space vector

$$\frac{d_c p_\iota}{dt} = \frac{dp_\iota}{dt} - \gamma_{\lambda,\iota\kappa}u^\kappa p^\lambda = \frac{dp_\iota}{dt} - \frac{1}{2}\frac{\partial\gamma_{\kappa\lambda}}{\partial x^\iota}u^\kappa p^\lambda, \tag{13}$$

the spatial covariant derivative of the space vector p_ι with respect to t. Here the $\gamma_{\lambda,\iota\kappa}$ are the three-dimensional Christoffel symbols formed by means of the spatial metric tensor $\gamma_{\iota\kappa}$, and the vector $d_c p_\iota/dt$ is thus the three-dimensional analogue of the four-vector (10*b*), (11).

Remembering that

$$\frac{dt}{d\tau} = \Gamma, \qquad \gamma_{\iota\kappa} = g_{\iota\kappa} + \gamma_\iota\gamma_\kappa, \tag{14}$$

it is easily seen that the equations (11) for $i = 1, 2, 3$ may be written in the form

$$\frac{d_c p_\iota}{dt} = K_\iota = mG_\iota, \tag{15}$$

where G_ι is a space vector depending on the dynamical gravitational potentials (γ_ι, χ) and their first derivatives. (See also Appendix 7.)

If $p^\iota = mu^\iota$ is interpreted as the momentum of the particle, K_ι must be interpreted as the (covariant) gravitational force acting on the particle. The proportionality factor m defined by (12) thus appears as the inertial mass of the particle moving with the velocity u^ι in the gravitational field with the potentials (γ_ι, χ). Since $K_\iota = mG_\iota$, m also represents the gravitational mass of the particle. For a particle at rest, the mass reduces to

$$m_0 = \frac{\mathring{m}_0}{\sqrt{(1+2\chi/c^2)}}, \tag{16}$$

which thus represents the rest mass of a particle in a gravitational field.

For small particle velocities, where we can neglect terms of the order u/c, the quantity G_ι is equal to a_ι given by (VIII. 95), but in general the gravitational force K_ι will be a complicated expression containing the potentials, their first derivatives, the velocity of the particle, and even its acceleration. There is one important case, however, where the expression for K_ι becomes extremely simple, viz. if the system of co-ordinates is time-orthogonal so that the vector potentials $\gamma_\iota = 0$. In this case it is immediately seen that the first three equations (11) are identical with (15) if we put

$$G_\iota = -\frac{\partial \chi}{\partial x^\iota}, \qquad \mathbf{K} = m\mathbf{G} = -m \operatorname{grad} \chi, \tag{17}$$

i.e. the gravitational force is connected with the scalar gravitational potential in the same way as in Newton's theory. This holds for arbitrarily strong fields and for all velocities (see also Appendix 7). Further, if $\gamma_\iota = 0$, the expression (12) for the 'relativistic' mass reduces to

$$m = \frac{\mathring{m}_0}{\sqrt{\left(1+\frac{2\chi}{c^2}-\frac{u^2}{c^2}\right)}}. \tag{17'}$$

On the other hand, for weak fields where the dynamical potentials (γ_ι, χ) may be treated as small, we find, neglecting terms of second order in these quantities as well as in u/c, that (11) is identical with (15) if we put

$$G_\iota = -\frac{\partial \chi}{\partial x^\iota} - c\frac{\partial \gamma_\iota}{\partial t} + c\omega_{\iota\kappa} u^\kappa, \tag{18}$$

where
$$\omega_{\iota\kappa} = \frac{\partial\gamma_\kappa}{\partial x^\iota} - \frac{\partial\gamma_\iota}{\partial x^\kappa}$$

is the space tensor defined by (VIII. 110) for the case of weak fields.

The last term in (18) is of the type of a Coriolis force. On the rotating disk considered in § 90, we have in the system of coordinates

$$x^i = (x, y, z, ct)$$

corresponding to the form (VIII. 83) of the line element, and in the approximation corresponding to (18)

$$\left.\begin{aligned} \gamma_\iota = g_{\iota 4} &= \frac{\omega}{c}(-y, x, 0) \\ \chi = \frac{c^2}{2}(-g_{44}-1) &= -\frac{r^2\omega^2}{2} \\ r^2 &= x^2+y^2 \end{aligned}\right\}. \tag{18'}$$

Hence
$$\omega_{\iota\kappa} = \begin{pmatrix} 0 & 2\frac{\omega}{c} & 0 \\ -2\frac{\omega}{c} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and for the gravitational force we get from (18)

$$K_\iota = (m\omega^2 x + 2m\omega u^2,\ m\omega^2 y - 2m\omega u^1,\ 0). \tag{18''}$$

For small distances r from the centre and for 'non-relativistic' particle velocities where terms of the order u^2/c^2 can be neglected, the gravitational force thus reduces to the usual combination of the centrifugal force and the Coriolis force.

Usually, the γ_ι (and their space derivatives) are smaller than or of the same order of magnitude as χ/c^2 (and its derivatives). For small values of u/c, the last term in (18) will then be small compared with the first term. Further, if the field is stationary or quasi-stationary, we may also neglect the second term in (18), and the gravitational force is again given by the Newtonian expression (17). Thus we see that, in the case of weak fields and small velocities, the contribution of the dynamical vector potential to the gravitational force is generally less important than that of the scalar potential. This is due to the large value of the constant c. The vector potential has an appreciable influence on the motion of a particle only for strongly fluctuating gravitational fields.

If also the geometrical influence of the gravitational field is weak, so that the physical space is approximately Euclidean, the motion of the particle is the same as that of a particle acted upon by a force of the type (17) in a system of inertia.

The spatial covariant derivative of the contravariant components p^ι with respect to t is defined by

$$\frac{d_c p^\iota}{dt} = \frac{dp^\iota}{dt} + \gamma^\iota_{\kappa\lambda} u^\kappa p^\lambda. \tag{13'}$$

Since the spatial metric tensor $\gamma_{\iota\kappa}$ may depend on t, $d_c p^\iota/dt$ and $d_c p_\iota/dt$ will not in general be the components of the same space vector. A calculation similar to that leading to (IX. 80‴) gives here

$$\frac{d_c p_\iota}{dt} = \gamma_{\iota\kappa}\frac{d_c p^\kappa}{dt} + \frac{\partial \gamma_{\iota\kappa}}{\partial t} p^\kappa \tag{19}$$

or the reciprocal relation

$$\frac{d_c p^\iota}{dt} = \gamma^{\iota\kappa}\frac{d_c p_\kappa}{dt} - \gamma^{\iota\kappa}\frac{\partial \gamma_{\kappa\lambda}}{\partial t} p^\lambda = K^\iota - \gamma^{\iota\kappa}\frac{\partial \gamma_{\kappa\lambda}}{\partial t} p^\lambda, \tag{19'}$$

where
$$K^\iota = \gamma^{\iota\kappa} K_\kappa$$

are the contravariant components of the gravitational force. The time derivative of the norm of the momentum vector is, by (19) and (19′),

$$\frac{d}{dt}(p_\iota p^\iota) = \frac{dp_\iota}{dt}p^\iota + p_\iota\frac{dp^\iota}{dt} = \frac{d_c p_\iota}{dt}p^\iota + p_\iota\frac{d_c p^\iota}{dt} = 2K_\iota p^\iota - \frac{\partial \gamma_{\iota\kappa}}{\partial t} p^\iota p^\kappa. \tag{20}$$

In a system where the dynamical potentials are zero, $\gamma_\iota = \chi = 0$, the gravitational force vanishes, and the motion of the particle is given by

$$\frac{d_c p_\iota}{dt} = 0, \tag{20'}$$

i.e. the covariant components of the momentum vectors at different times are obtained by parallel displacements in the three-dimensional sense. In general this does not mean, however, that the magnitude of the momentum vector is constant in time, for from (20) we get

$$\frac{dp^2}{dt} = -\frac{\partial \gamma_{\iota\kappa}}{\partial t} p^\iota p^\kappa. \tag{20''}$$

Thus, p is constant only if our system of reference is rigid, i.e. if $\gamma_{\iota\kappa}$ is time-independent. If the dynamical potentials are zero, we have now

$$m = \frac{\mathring{m}_0}{\sqrt{(1-u^2/c^2)}}, \qquad p^2 = m^2u^2 = \frac{\mathring{m}_0^2 u^2}{1-u^2/c^2},$$

and, if further the frame of reference is rigid, both u and m are constant and the equations of motion (20′) may be written

$$\frac{d_c u_\iota}{dt} = 0. \tag{20'''}$$

Hence, in this case the orbit of the particle is a geodesic line in physical space, i.e. the particle is moving with constant velocity in the 'straightest' line compatible with the geometry of the space. The motion of the particle is thus completely analogous to the motion of a free particle bound to move on a smooth curved two-dimensional surface in a system of inertia where the only forces on the particle are the normal reactions of the surface. The only essential difference is that here we have to deal with the motion of a particle in a curved three-dimensional space.

If the spatial metric tensor varies with t, the motion of the particle in the gravitational field is analogous to the motion of a particle on a smooth *variable* surface in a system of inertia. Thus, if the dynamical potentials are zero, the action of the gravitational field has the character of a 'normal reaction' from the curved three-dimensional space.

111. Total energy of a particle in a stationary gravitational field

While the first three equations of the set (11) represent the equations of motion of the particle, the fourth equation corresponding to $i = 4$ must be the law of conservation of energy. In this section, we consider only the case of stationary fields, where we have no flux of gravitational energy, and leave the general case to a later section (§ 126). Hence the g_{ik} are time-independent and the equation in question becomes

$$\frac{dP_4}{d\tau} = 0, \tag{21}$$

i.e. P_4 is a constant of the motion in this case. The constant $H = -cP_4$, which on account of (4) is of the form

$$\begin{aligned} H = -cP_4 &= \mathring{m}_0 c^2 \Gamma\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\} \\ &= \frac{\mathring{m}_0 c^2\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}}{\left[\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}^2 - \frac{u^2}{c^2}\right]^{\frac{1}{2}}}, \end{aligned} \tag{22}$$

may be interpreted as the total energy of the particle in the gravitational field (see Appendix 7). For small velocities we get, neglecting terms of *second* and higher order in u/c,

$$H_0 = \mathring{m}_0 c^2\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}, \tag{23}$$

since the terms of first order in the velocities cancel. H_0 represents the rest energy of a particle in the field. If we retain the terms of second order in u/c, we get, using (16),

$$H = H_0 + \tfrac{1}{2} m_0 u^2, \tag{24}$$

and for weak fields this reduces to the usual expression for the energy of a particle in a field with the gravitational potential χ, i.e.

$$H = \mathring{m}_0 c^2 + \tfrac{1}{2}\mathring{m}_0 u^2 + \mathring{m}_0 \chi. \tag{25}$$

The last two terms represent the kinetic energy and the potential energy, respectively. For large velocities such a decomposition of the energy into a kinetic part and a potential part is not possible.

112. General point mechanics

If the particle is acted upon by non-gravitational forces also, for instance electromagnetic forces, we have, on account of (IV. 55, 57), in a general system of coordinates

$$\frac{DP_\iota}{d\tau} = \frac{dP_\iota}{d\tau} - \frac{1}{2}\frac{\partial g_{kl}}{\partial x^\iota} U^k P^l = F_\iota, \tag{26}$$

$$F_\iota U^\iota = 0, \tag{27}$$

where F_ι is a four-vector which in a local system of inertia is identical with Minkowski's four-force (IV. 54). In the first instance, the physical meaning of the quantities F_ι may thus be obtained by a transformation of the physically well-defined Minkowski four-force. The components F_ι may, however, also be given a simple physical interpretation in the system of coordinates (x^ι) itself. If we put $F_\iota = \Gamma \mathfrak{F}_\iota$, we get from the equations (27) and (2)

$$F_4 = -\frac{1}{c}\Gamma(\mathfrak{F}_\iota u^\iota) = -\frac{\Gamma}{c}(\mathbf{\mathfrak{F}} \cdot \mathbf{u}),$$

i.e.

$$F_\iota = \left(\Gamma \mathbf{\mathfrak{F}},\ -\frac{\Gamma}{c}(\mathbf{\mathfrak{F}} \cdot \mathbf{u})\right). \tag{28}$$

Since $dt = d\tau . \Gamma$, we find, in the same way as in § 110, that the first three equations (26) may be written

$$\frac{d_c p_\iota}{dt} = K_\iota + \mathfrak{F}_\iota, \tag{29}$$

which shows that the space vector $\mathbf{\mathfrak{F}}$ with the covariant components $\mathfrak{F}_\iota = F_\iota/\Gamma$ must be interpreted as the non-gravitational force on the particle.

Further, in the case of a stationary gravitational field, the fourth equation (26) can be written

$$\frac{dP_4}{d\tau} = F_4,$$

or, by means of (14), (22), and (28),

$$\frac{dH}{dt} = (\mathfrak{F}.\mathbf{u}) = \mathfrak{F}_\iota u^\iota. \tag{30}$$

Since the right-hand side is equal to the work done by the force $\mathfrak{F}$ per unit time, this equation expresses the law of conservation of energy.

113. Time-orthogonal systems of coordinates. Elimination of the dynamical potentials

As shown in § 94, it is not in general possible by a transformation of the type (VIII. 59) (i.e. without a change of the system of reference), to ensure that the transformed vector potentials γ'_ι or $g'_{\iota 4}$ vanish. The condition for this to be possible is that

$$\omega_{\iota\kappa} = 0,$$

where $\omega_{\iota\kappa}$ is the space tensor defined by (VIII. 110).

However, if we allow arbitrary changes of the system of reference, i.e. by means of a general space-time transformation

$$x'^\iota = f^\iota(x^k), \tag{31}$$

we can always obtain

$$g'_{\iota 4} = 0 \tag{32}$$

even for the most general type of gravitational fields described by arbitrary functions $g_{\iota k}$ or $g^{\iota k}$ in the system (x^i).

From (IX. 7) we get, if (32) is satisfied,

$$g'^{\iota l} g'_{4l} = g'^{\iota 4} g'_{44} = \delta^\iota_4,$$

i.e.

$$g'^{\iota 4} = 0, \qquad g'^{44} = \frac{1}{g'_{44}}. \tag{33}$$

The transformation law of a tensor

$$g'^{\iota k} = \frac{\partial f^\iota}{\partial x^l}\frac{\partial f^k}{\partial x^m} g^{lm} \tag{34}$$

in connexion with the condition $g'^{\iota 4} = 0$ thus leads to the equations

$$0 = \frac{\partial f^\iota}{\partial x^l}\frac{\partial f^4}{\partial x^m} g^{lm} \qquad (\iota = 1, 2, 3). \tag{35}$$

For $f^4(x^k)$ we can now choose an arbitrary function satisfying the necessary condition (VIII. 50):

$$g'^{44} = \frac{1}{g'_{44}} = \frac{\partial f^4}{\partial x^l}\frac{\partial f^4}{\partial x^m} g^{lm} = \operatorname{grad}_i f^4 \operatorname{grad}^\iota f^4 < 0. \tag{36}$$

For the functions $f^{\iota}(x)$ we then get the three linear partial differential equations of first order

$$\mathrm{grad}^{l} f^{4} \frac{\partial f^{\iota}}{\partial x^{l}} = 0. \tag{37}$$

Each of the functions f^{ι} thus must satisfy a partial differential equation of the type

$$A^{l} \frac{\partial f}{\partial x^{l}} = 0, \tag{38}$$

where the quantities $$A^{l} = \mathrm{grad}^{l} f^{4}$$

may be considered known functions.

The general solution of an equation of this type contains an arbitrary function of three variables which may be three of the independent variables or arbitrary combinations of them. For the functions f^{ι}, we can thus take three arbitrary independent solutions of (38), and the transformations (31) will then lead to a system of coordinates in which $g'^{\iota 4} = g'_{\iota 4} = 0$ or $\gamma'_{\iota} = 0$. In this way the vector potential has been 'transformed away' and we have obtained a time-orthogonal system of coordinates. On account of the simplifications which arise from the disappearance of the vector potentials, in the future we often take advantage of this possibility and use time-orthogonal systems of coordinates.

In the preceding discussion, the function $f^{4}(x)$ was chosen arbitrarily apart from the very mild restrictive condition (36). We shall now see that it is also possible, by a suitable choice of the function $f^{4}(x)$, to make the scalar potential vanish. On accouht of (34), the condition $\chi' = 0$, i.e. $g'^{44} = 1/g'^{44} = -1$ leads to the condition

$$-1 = \frac{\partial f^{4}}{\partial x^{l}} \frac{\partial f^{4}}{\partial x^{m}} g^{lm} = \mathrm{grad}_{\iota} f^{4} \, \mathrm{grad}^{\iota} f^{4} \tag{39}$$

for the function f^{4}, i.e. the norm of the gradient of the function $f^{4}(x)$ must be constant and equal to -1. Geometrically this means that the manifold of hypersurfaces defined by $f^{4}(x) =$ constant are at a constant distance from each other. This can obviously be obtained in an infinite number of ways, since we can choose one space-like surface arbitrarily and construct all the consecutive surfaces such that the distance between two surfaces measured along the normal is constant for all points on the surface. Now, substituting the function f^{4} obtained in this way in (37), and solving these equations, we get a set of transformation functions $f^{i}(x)$ which leads to a system of coordinates in which the four dynamical gravitational potentials are zero. In a system of this type, the gravitational force $\mathbf{K}$ occurring in (15) is zero, and the motion of a free particle is of the type discussed on p. 293. If the initial velocity of the particle is zero, we see at once from (20″) that the particle will remain at rest in this system. The different points of reference in the system may thus be represented by an assembly of freely falling material particles, a circumstance which may be used in the practical determination of the transformation functions $f^{\iota}(x)$.

Although a permanent gravitational field cannot be completely transformed away except in infinitesimal regions of space-time, it is thus always possible to eliminate the *dynamical* effects of a gravitational field over finite regions of 4-space. The effect of the gravitational field is then purely geometrical and is completely described by the spatial metric tensor. Although very interesting from the theoretical point of view, the possibility of transforming away the dynamical properties of the gravitational fields is usually of little practical importance, since the corresponding system of reference is generally not rigid and the time-dependence of the spatial metric tensor makes a treatment of physical phenomena in this system very complicated. In the treatment of cosmological problems in Chapter XII we shall, however, make use of this possibility. (See §§ 133, 134.)

114. Mechanics of continuous systems

By the same procedure as that applied in §§ 109–12, all physical laws of the special theory of relativity may now easily be written in a general covariant form. In Chapter VI it was shown that the behaviour of a closed system in the special theory could be described by a symmetrical energy-momentum tensor $T_{\iota k}$ satisfying the equation (VI. 1) which in the real representation may be written in the form

$$\operatorname{div}_{\iota}\{T^{\iota k}\} = \frac{\partial T_{\iota}^{k}}{\partial x^{k}} = 0. \tag{40}$$

The physical meaning of the different components was explained in § 62. Since the equation (40) is assumed to hold in any local system of inertia, the general covariant form of the laws of conservation of energy and momentum must be (see IX. 92c)

$$\left.\begin{aligned} T^{k}_{\iota,k} &= 0 \\ \text{or}\qquad \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^{k}}(\sqrt{(|g|)}T_{\iota}^{k}) &= \frac{1}{2}\frac{\partial g_{kl}}{\partial x^{\iota}}T^{kl} \end{aligned}\right\}. \tag{41}$$

While the conservation laws of a closed system in a system of inertia are expressed by an equation containing a sum of partial derivatives, we see that this is not so in the general case on account of the term on the right-hand side of (41). This indicates that the system is no longer closed when placed in a gravitational field which itself may contribute to the total energy and momentum. We shall return to this question in § 126.

The components of the tensor T_i^k in an arbitrary system of coordinates S may be obtained from the tensor $\overset{0}{T}{}_i^k$ in a local system of inertia $\overset{0}{S}$ by means of the law of transformation of tensors and, since the physical meaning of $\overset{0}{T}{}_i^k$ is the same as in the special theory of relativity, we get, in this way, a physical interpretation of the T_i^k also. In $\overset{0}{S}$ we can apply all the considerations of §§ 63, 64, i.e. if the physical system is so small that ϑ_{ik} may be regarded as constant over the whole region occupied by the system, we can unambiguously define the proper centre of mass of the system. Relative to $\overset{0}{S}$ this point is moving with constant velocity and its motion relative to S will thus be like the motion of a freely falling particle †

The three first equations (41) represent the law of conservation of momentum and the fourth equation is the law of conservation of energy. If the gravitational field is stationary, the right-hand side of the equation (41) with $i = 4$ is zero. Hence

$$\frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^k}(\sqrt{(|g|)}T_4^k) = 0. \tag{42}$$

Appendix 8 shows that the determinant $\gamma = |\gamma_{\iota\kappa}|$, formed by the components of the spatial metric tensor (VIII. 64), is connected with the determinant $g = |g_{ik}|$ by the equation $g = g_{44}\cdot\gamma$ and, since $g < 0$, $\gamma > 0$, we have

$$|g| = -g_{44}\cdot\gamma = \gamma\left(1+\frac{2\chi}{c^2}\right). \tag{43}$$

The equation (42) may thus be written

$$\frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial x^k}\left\{\sqrt{\gamma}\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} T_4^k\right\} = 0. \tag{44}$$

Now, putting

$$S^\iota = -c\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} T_4^\iota, \qquad h = -\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} T_4^4 \tag{45}$$

† A. D. Fokker, *Proc. Amsterdam*, **23**, No. 5, 729 (1921). C. Møller, *Annales de l'Institut Henri Poincaré*, t. XI, fasc. V, 251 (1950).

and remembering that $\sqrt{\gamma}$ is time-independent in the case considered, (44) may be written

$$\operatorname{div}\mathbf{S}+\frac{\partial h}{\partial t}=0, \tag{46}$$

where

$$\operatorname{div}\mathbf{S}=\frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial x^{\iota}}(\sqrt{\gamma}S^{\iota})$$

is the three-dimensional divergence of the space vector **S** (cf. the analogous expression (IX. 90) for the four-dimensional divergence).

The equation (46) expresses the law of conservation of energy if h and **S** are interpreted as the energy density and energy flux, respectively. In a local system of inertia, (45) is identical with the equations (VI. 2, 3). As we shall see in § 126, the momentum density is similarly given by

$$g_{\iota}=\frac{1}{c}\Big(1+\frac{2\chi}{c^2}\Big)^{\frac{1}{2}}T_{\iota}^{4}. \tag{47}$$

If the physical system is a perfect fluid, the energy-momentum tensor has the form

$$T_{\iota}^{k}=\Big(\overset{0}{\mu}{}^{0}+\frac{\overset{0}{p}}{c^2}\Big)U_{\iota}U^{k}+\overset{0}{p}\,\delta_{\iota}^{k}, \tag{48}$$

where $U^{\iota}=U^{\iota}(x^k)$ is the four-velocity of the matter at the event point (x^k), while $\overset{0}{\mu}{}^{0}$ and $\overset{0}{p}$ represent the invariant mass density and pressure, respectively, measured in a local rest system of inertia. The expression (48) transforms like a mixed tensor and, since it reduces to the expression (VI. 104) in a local system of inertia, it must be the correct expression for T_{ι}^{k} in every system of coordinates.

For incoherent matter we have simply

$$T_{\iota}^{k}=\overset{0}{\mu}{}^{0}U_{\iota}U^{k}. \tag{49}$$

By means of the general rule exemplified in (IX. 84, 85), the conservation laws (41) can be written

$$(\overset{0}{\mu}{}^{0}U^{k})_{,k}.U_{\iota}+(\overset{0}{\mu}{}^{0}U^{k}).U_{\iota,k}=0. \tag{50}$$

From (6) we get by covariant differentiation

$$U_{\iota,k}U^{\iota}+U_{\iota}.U^{\iota}{}_{,k}=0,$$

and since

$$U_{\iota}U^{\iota}{}_{,k}=U^{\iota}.U_{\iota,k},$$

we have

$$U^{\iota}.U_{\iota,k}=0. \tag{51}$$

Thus, if we multiply (50) by U^{ι} we obtain

$$(\overset{0}{\mu}{}^{0}U^{k})_{,k}=\operatorname{div}\{\overset{0}{\mu}{}^{0}U^{k}\}=\frac{1}{\sqrt{|g|}}\frac{\partial(\sqrt{|g|}\overset{0}{\mu}{}^{0}U^{k})}{\partial x^{k}}=0. \tag{52}$$

This equation, which is the generalization of the equation (IV. 211), expresses the conservation of proper mass.

On account of (IX. 79, 80″) we now have

$$U^k U_{i;k} = \frac{dx^k}{d\tau}\left(\frac{\partial U_i}{\partial x^k} - \Gamma^l_{ik}\, U_l\right) = \frac{dU_i}{d\tau} - \Gamma^l_{ik}\, U^k U_l = \frac{DU_i}{d\tau},$$

and (50) thus reduces to

$$\overset{0}{\mu}{}^0 . \frac{DU_i}{d\tau} = 0, \tag{53}$$

which is the generalization of (IV. 215) for the case in which the gravitational forces are the only forces acting on the matter. A comparison with (IX. 80^{IV}) shows that each infinitesimal part of incoherent continuously distributed matter is moving like a freely falling particle.

From (45), (49), (2), and (4) we get for the energy density

$$h = -\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} \overset{0}{\mu}{}^0 U_4\, U^4 = \overset{0}{\mu}{}^0 c^2 \Gamma^2 \left(1+\frac{2\chi}{c^2}\right)\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}, \tag{54}$$

where Γ is given by (3). Now consider a small piece of matter with the volume $d\overset{0}{V}{}^0 = d\overset{0}{x}{}^1 d\overset{0}{x}{}^2 d\overset{0}{x}{}^3$ in a local rest system of inertia $\overset{0}{S}{}^0$. On account of the invariance of the four-dimensional volume element (IX. 47) we have, since the proper time τ is identical with the time in $\overset{0}{S}{}^0$,

$$c\sqrt{(|g|)}\, dx^1 dx^2 dx^3 dt = c\, d\overset{0}{x}{}^1 d\overset{0}{x}{}^2 d\overset{0}{x}{}^3 d\tau.$$

By means of (43) and (14) we thus get for the volume dV of the material particle

$$\left.\begin{aligned} dV &= \sqrt{\gamma}\, dx^1 dx^2 dx^3 = \frac{d\overset{0}{V}{}^0}{\sqrt{(1+2\chi/c^2)}}\frac{d\tau}{dt} = \frac{d\overset{0}{V}{}^0}{\Gamma\sqrt{(1+2\chi/c^2)}} \\ dV &= \frac{d\overset{0}{V}{}^0\left[\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\}^2 - \frac{u^2}{c^2}\right]^{\frac{1}{2}}}{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}} \end{aligned}\right\}. \tag{55}$$

This is the generalization of the Lorentz contraction formulae in the presence of gravitational fields. For $u^\kappa = 0$, we get

$$dV^0 = d\overset{0}{V}{}^0, \tag{56}$$

i.e. the volume of a small piece of matter at rest in S is equal to its volume measured in $\overset{0}{S}{}^0$ in accordance with the general assumption made on p. 223.

If $\overset{0}{m}_0 = \overset{0}{\mu}{}^0\, d\overset{0}{V}{}^0$ denotes the proper mass of the particle, its total energy $H = h\, dV$ is obtained from (54) and (55)

$$H = h\, dV = \overset{0}{m}_0 c^2 \Gamma\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\left\{\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} - \frac{\gamma_\kappa u^\kappa}{c}\right\} \tag{57}$$

in accordance with (22).

115. The electromagnetic field equations

For simplicity we confine ourselves to the case of electrodynamics in the vacuum, the generalization of the macroscopic theory in ponderable matter running along the same lines. By means of the generally covariant expressions for the divergence and curl of an antisymmetrical tensor (IX. 93, 94), we can write Maxwell's equations (V. 9, 13, 16) in an arbitrary system of coordinates S in the form

$$\frac{\partial F_{ik}}{\partial x^l}+\frac{\partial F_{kl}}{\partial x^i}+\frac{\partial F_{li}}{\partial x^k}=0, \tag{58a}$$

$$\frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^k}(\sqrt{|g|}\,F^{ik})=s^i, \tag{58b}$$

$$s^i=\frac{\overset{0}{\rho}U^i}{c}. \tag{59}$$

Here, $F^{ik}=-F^{ki}$ is the electromagnetic field tensor, s^i is the four-current density, $\overset{0}{\rho}$ is the charge density measured in a local rest system of inertia, and U^i is the four-velocity of the electric charge. In every local system of inertia $\overset{0}{S}$, the equations (58) and (59) reduce to the equations (13), (16), or (1) in Chapter V.

On account of the antisymmetry of F^{ik} we get at once from (58 b)

$$\operatorname{div}\{s^i\}=\frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^i}(\sqrt{|g|}s^i)=\frac{1}{\sqrt{|g|}}\frac{\partial^2}{\partial x^i\partial x^k}(\sqrt{|g|}F^{ik})\equiv 0, \tag{60}$$

which is the general form of the continuity equation of electric charge.

Now consider a system which at the time $t=x^4/c$ lies entirely inside a finite region of volume V in 3-space. Multiplying (60) by $\sqrt{|g|}$ and integrating over the space coordinates x^1, x^2, x^3 we get, since the first three terms in the summation over i are partial differential coefficients with respect to x^1, x^2, x^3,

$$\frac{d}{dx^4}\int\limits_V\sqrt{|g|}\,s^4\,dx^1dx^2dx^3=0. \tag{61}$$

Thus the quantity

$$e=\int\limits_V\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}s^4\sqrt{\gamma}\,dx^1dx^2dx^3=\int\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}s^4\,dV \tag{62}$$

is a constant in time, which must be interpreted as the total electric charge of the system. Hence

$$\rho=\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}s^4=\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}}\Gamma\overset{0}{\rho} \tag{63}$$

must be the charge density in the system S. Now consider the charge $\rho\, dV$ connected with a volume element dV; if $d\overset{0}{V}{}^0$ represents the corresponding volume in the local rest system, we get from (63) and (55)

$$\rho\, dV = \overset{0}{\rho}{}^0\, d\overset{0}{V}{}^0, \tag{64}$$

i.e. the electric charge of the small piece of matter considered is independent of the system of coordinates.

The invariance of the *total* electric charge also follows directly from the continuity equation (60). Consider two arbitrary systems of coordinates S and S' and the integrals

$$e = \int_{x^4=a} \sqrt{|g|}\, s^4\, dx^1 dx^2 dx^3 \tag{65}$$

$$e' = \int_{x'^4=b} \sqrt{|g'|}\, s'^4\, dx'^1 dx'^2 dx'^3 \tag{65'}$$

over the two regions defined by

$$x^4 = \text{constant} = a \quad \text{and} \quad x'^4 = \text{constant} = b,$$

respectively. On account of the time-independence of the integrals (65) and (65′) we can, without any change in the values of e and e', choose b so that the regions $x'^4 = b$ and $x^4 = a$ do not overlap inside the tube in 4-space in which the charge density is different from zero. Then we can introduce a third system of coordinates S'' which coincides with S inside the region $x^4 = a$ and with S' inside the region $x'^4 = b$. As the equation (61) holds also in the system S'', we have

$$\int_{x''^4=x^4=a} \sqrt{|g''|}\, s''^4\, dx''^1 dx''^2 dx''^3 = \int_{x''^4=x'^4=b} \sqrt{|g''|}\, s''^4\, dx''^1 dx''^2 dx''^3 \tag{66}$$

and, since $x''^\iota = x^\iota$ on the hypersurface $x^4 = a$ and $x''^\iota = x'^\iota$ in the region $x'^4 = b$, (66) leads to the equation

$$e = e', \tag{67}$$

which expresses the invariance of the total charge.

On account of (59), (63), and (2) we have

$$s^\iota = \left(\frac{\rho u^\iota}{c\sqrt{(1+2\chi/c^2)}}, \frac{\rho}{\sqrt{(1+2\chi/c^2)}}\right), \tag{68}$$

and the continuity equation (60) can be written

$$\frac{1}{\sqrt{\gamma}} \frac{\partial}{\partial x^\iota}(\sqrt{\gamma}\rho u^\iota) + \frac{1}{\sqrt{\gamma}} \frac{\partial \sqrt{\gamma}\rho}{\partial t} = 0 \quad \text{or} \quad \operatorname{div} \rho \mathbf{u} + \frac{1}{\sqrt{\gamma}} \frac{\partial \sqrt{\gamma}\rho}{\partial t} = 0, \tag{69}$$

which is the three-dimensional form of the continuity equation when the spatial metric is time-dependent. For rigid frames of reference, where γ does not depend on t, (69) reduces to the usual form

$$\operatorname{div}\rho\mathbf{u}+\frac{\partial\rho}{\partial t}=0.$$

In order to obtain a closer understanding of the physical meaning of the tensor components F^{ik} we shall now write the Maxwell equations (58) in the form of three-dimensional vector equations. For simplicity we assume that our system of coordinates is time-orthogonal, i.e.

$$\gamma_\iota = g_{\iota 4} = 0, \qquad \gamma_{\iota\kappa} = g_{\iota\kappa}, \qquad -g_{44} = 1+\frac{2\chi}{c^2}, \tag{70}$$

which, as was shown in § 113, can always be obtained by a suitable choice of the space-time coordinates.

Let us now introduce two antisymmetrical space tensors $H^{\iota\kappa}$ and $B^{\iota\kappa}$ and two space vectors D^ι and E^ι by the equations

$$F^{\iota\kappa} = \frac{H^{\iota\kappa}}{\sqrt{(1+2\chi/c^2)}} = B^{\iota\kappa}, \qquad F^{\iota 4} = -\frac{D^\iota}{\sqrt{(1+2\chi/c^2)}} = -\frac{E^\iota}{1+2\chi/c^2}. \tag{71}$$

These quantities are thus connected by the equations

$$B^{\iota\kappa} = \frac{H^{\iota\kappa}}{\sqrt{(1+2\chi/c^2)}}, \qquad \mathbf{D} = \frac{\mathbf{E}}{\sqrt{(1+2\chi/c^2)}}. \tag{72}$$

From (IX. 27) and the corresponding equations for spatial tensors we then get, using (70) and (71),

$$\left.\begin{aligned} F_{\iota\kappa} &= g_{\iota l}g_{\kappa m}F^{lm} = g_{\iota\lambda}g_{\kappa\mu}F^{\lambda\mu} = \gamma_{\iota\lambda}\gamma_{\kappa\mu}B^{\lambda\mu} = B_{\iota\kappa} = \frac{H_{\iota\kappa}}{\sqrt{(1+2\chi/c^2)}} \\ F_{\iota 4} &= g_{\iota l}g_{4m}F^{lm} = \gamma_{\iota\kappa}\left(1+\frac{2\chi}{c^2}\right)\frac{D^\kappa}{\sqrt{(1+2\chi/c^2)}} = D_\iota\left(1+\frac{2\chi}{c^2}\right)^{\frac{1}{2}} = E_\iota \end{aligned}\right\}. \tag{73}$$

With these notations, the equations (58 a), (58 b) may be written in the form of three-dimensional tensor equations:

$$\left.\begin{aligned} &\operatorname{curl}_{\iota\kappa\lambda}\{B_{\iota\kappa}\} = 0, \qquad \operatorname{curl}_{\iota\kappa}\mathbf{E} = -\frac{1}{c}\frac{\partial B_{\iota\kappa}}{\partial t} \\ &\operatorname{div}^\iota\{H^{\iota\kappa}\} - \frac{1}{\sqrt{\gamma}}\frac{1}{c}\frac{\partial(\sqrt{\gamma}D^\iota)}{\partial t} = \frac{\rho u^\iota}{c}, \qquad \operatorname{div}\mathbf{D} = \rho \end{aligned}\right\}, \tag{74}$$

where

$$\left.\begin{aligned} \operatorname{curl}_{\iota\kappa\lambda}\{B_{\iota\kappa}\} &= \frac{\partial B_{\iota\kappa}}{\partial x^\lambda}+\frac{\partial B_{\kappa\lambda}}{\partial x^\iota}+\frac{\partial B_{\lambda\iota}}{\partial x^\kappa} \\ \operatorname{curl}_{\iota\kappa}\mathbf{E} &= \frac{\partial E_\kappa}{\partial x^\iota}-\frac{\partial E_\iota}{\partial x^\kappa} \\ \operatorname{div}^\iota\{H^{\iota\kappa}\} &= \frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial x^\kappa}(\sqrt{\gamma}H^{\iota\kappa}) \\ \operatorname{div}\mathbf{D} &= \frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial x^\iota}(\sqrt{\gamma}D^\iota) \end{aligned}\right\} \tag{75}$$

are the three-dimensional covariant differential operators. Further, defining the axial vectors dual to the antisymmetrical tensors $B_{\iota\kappa}$, $H^{\iota\kappa}$ by

$$\left.\begin{aligned} B^1 &= \frac{1}{\sqrt{\gamma}}B_{23}, & B^2 &= \frac{1}{\sqrt{\gamma}}B_{31}, & B^3 &= \frac{1}{\sqrt{\gamma}}B_{12} \\ H_1 &= \sqrt{\gamma}H^{23}, & H_2 &= \sqrt{\gamma}H^{31}, & H_3 &= \sqrt{\gamma}H^{12} \end{aligned}\right\} \tag{76}$$

(see Appendix 5), Maxwell's equations (74) take the familiar form

$$\left.\begin{aligned} \operatorname{curl}\mathbf{E} &= -\frac{1}{c}\frac{1}{\sqrt{\gamma}}\frac{\partial(\sqrt{\gamma}\mathbf{B})}{\partial t}, & \operatorname{div}\mathbf{B} &= 0 \\ \operatorname{curl}\mathbf{H}-\frac{1}{\sqrt{\gamma}}\frac{\partial(\sqrt{\gamma}\mathbf{D})}{c\,\partial t} &= \frac{\rho\mathbf{u}}{c}, & \operatorname{div}\mathbf{D} &= \rho \end{aligned}\right\}, \tag{77}$$

where curl **E** and curl **H** are the vectors dual to the tensors $\operatorname{curl}_{\iota\kappa}\mathbf{E}$ and $\operatorname{curl}_{\iota\kappa}\mathbf{H}$, respectively. For a rigid system of reference, where γ is time-independent, these vector equations are of the same form as Maxwell's phenomenological equations in ponderable matter and, since (72) can be written

$$\mathbf{D} = \epsilon\mathbf{E}, \qquad \mathbf{B} = \mu\mathbf{H}, \tag{78a}$$

$$\epsilon = \mu = \frac{1}{\sqrt{(1+2\chi/c^2)}}, \tag{78b}$$

we see that the gravitational field, besides its influence on the spatial geometry, acts like a medium with the dielectric and magnetic constants (78·*b*).

116. Electromagnetic force and energy-momentum tensor

The electromagnetic four-force acting on a particle with charge e is, according to (V. 85),

$$F_i = \frac{e}{c}F_{ik}U^k. \tag{79}$$

For the components of this vector we get, by means of (2) and (73),

$$F_i = \frac{e}{c}(F_{\iota\kappa}U^\kappa+F_{\iota 4}U^4,\ F_{4\kappa}U^\kappa) = e\Gamma\left(E_\iota+\frac{B_{\iota\kappa}u^\kappa}{c},\ -\frac{E_\kappa u^\kappa}{c}\right).$$

A comparison with (28) shows that the electromagnetic force on a charged moving particle is

$$\mathfrak{F}_\iota = e\left(E_\iota + \frac{B_{\iota\kappa} u^\kappa}{c}\right), \tag{80}$$

which, by means of the axial vector **B** defined by (76), may be written

$$\mathfrak{F} = e\left(\mathbf{E} + \frac{\mathbf{u}\times\mathbf{B}}{c}\right), \tag{81}$$

where $\mathbf{u}\times\mathbf{B}$ is the vector product with the covariant components

$$(\mathbf{u}\times\mathbf{B})_\iota = \sqrt{\gamma}(u^2B^3-u^3B^2,\ u^3B^1-u^1B^3,\ u^1B^2-u^2B^1) \tag{81'}$$

(see Appendix 5).

For a system with continuously distributed charge the force on the charge in a volume element dV is given by (81) with $e = \rho\, dV$. Therefore the force *density* **f** must be

$$\mathbf{f} = \rho\left(\mathbf{E} + \frac{\mathbf{u}\times\mathbf{B}}{c}\right), \tag{82}$$

which is formally identical with the expression for the ponderomotive force (VII. 73) following from Minkowski's energy-momentum tensor in ponderable matter.

For the four-force density we get from (V. 93)

$$f_\iota = F_{\iota k} s^k, \tag{83}$$

which, on account of (68) and (73), leads to the expressions

$$f_\iota = \left(\frac{\rho(E_\iota + B_{\iota\kappa} u^\kappa/c)}{\sqrt{(1+2\chi/c^2)}},\ -\frac{\rho(E_\kappa u^\kappa)}{c\sqrt{(1+2\chi/c^2)}}\right). \tag{84}$$

Hence
$$f_\iota = \left(\frac{\mathbf{f}}{\sqrt{(1+2\chi/c^2)}},\ -\frac{(\mathbf{f}.\mathbf{u})}{c\sqrt{(1+2\chi/c^2)}}\right), \tag{85}$$

which represents the generalization of (IV. 214) in the presence of a gravitational field.

According to (V. 96) the equations of motion, in the general system S, are

$$\mathring{\mu}^0 \frac{DU_\iota}{d\tau} = f_\iota. \tag{86}$$

If we multiply this equation by the proper volume dV^0 of a small piece of charged matter, then, since the proper mass $\mathring{\mu}^0\, d\mathring{V}^0$ is conserved by this type of force, we get, using (55),

$$\frac{DP_\iota}{d\tau} = f_\iota\, d\mathring{V}^0 = f_\iota\, dV\, \Gamma\sqrt{\left(1+\frac{2\chi}{c^2}\right)}, \tag{87}$$

where $P_\iota = \overset{0}{\mu}{}^0\, d\overset{0}{V}{}^0 U_\iota$ is the four-momentum of the small particle. On account of (85) and (14), these equations may also be written

$$\frac{d_c \mathbf{p}}{dt} = \mathbf{K} + \mathbf{f}\, dV, \qquad \frac{dH}{dt} = (\mathbf{f}.\mathbf{u})\, dV, \tag{88}$$

similar to (29) and (30).

The first equations are the equations of motion of the small piece of matter under the influence of the gravitational force $\mathbf{K}$ and the electromagnetic force $\mathbf{f}\, dV$, the last equation is the law of conservation of energy.

From the validity of the equations (V. 105, 106) in a local system of inertia we get in the general system S

$$f_\iota = -\mathrm{div}_\iota\{S^k_\iota\} \equiv S^k_{\iota,k} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^k}(\sqrt{|g|} S^k_\iota) - \frac{1}{2} \frac{\partial g_{kl}}{\partial x^\iota} S^{kl}, \tag{89}$$

where
$$S^k_\iota = S_\iota^{\;k} = S^{k}_{\cdot\,\iota} = F_{\iota l} F^{kl} - \delta^k_\iota \tfrac{1}{4}(F_{lm} F^{lm}) \tag{90}$$

is the electromagnetic energy-momentum tensor. The equation (89) is a consequence of the field equations (58) together with (83).

Using (71), (73), and (76) in (90) we get for the different components of the tensor S^k_ι

$$S^\kappa_\iota = -\frac{t^\kappa_\iota}{\sqrt{(1+2\chi/c^2)}}, \qquad t^\kappa_\iota = H_\iota B^\kappa + E_\iota D^\kappa - \delta^\kappa_\iota \tfrac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B}), \tag{91 a}$$

$$S^\iota_4 = -\frac{(\mathbf{E}\times\mathbf{H})^\iota}{\sqrt{(1+2\chi/c^2)}}, \qquad S^4_4 = -\frac{\tfrac{1}{2}(\mathbf{E}\;\mathbf{D}+\mathbf{H}.\mathbf{B})}{\sqrt{(1+2\chi/c^2)}}, \tag{91 b}$$

$$S^4_\iota = \frac{(\mathbf{D}\times\mathbf{B})_\iota}{\sqrt{(1+2\chi/c^2)}}, \tag{91 c}$$

where $\mathbf{E}\times\mathbf{H}$ and $\mathbf{D}\times\mathbf{B}$ are the vector products defined by (81′) or by the equations (5) and (6) in Appendix 5.

From (45), (47), and (91) we get for the electromagnetic energy flux, energy density, and momentum density

$$\mathbf{S} = c(\mathbf{E}\times\mathbf{H}), \qquad h = \tfrac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B}), \tag{92}$$

$$\mathbf{g} = \frac{1}{c}(\mathbf{D}\times\mathbf{B}). \tag{93}$$

These expressions are in complete agreement with the formulae (VII. 70, 71, 72) following from Minkowski's energy-momentum tensor (VII. 68) in ponderable matter, as distinct from Abraham's tensor which gives $\mathbf{g} = (\mathbf{E}\times\mathbf{H})/c$ instead of (93).

117. Propagation of light in a static gravitational field. Fermat's principle

In a static gravitational field, where $\gamma_\iota = 0$ and $\gamma_{\iota\kappa}$ and χ are independent of the time variable t, the electromagnetic field equations (77) take the form of Maxwell's phenomenological equations in a medium at rest with

$$\epsilon = \mu = \frac{1}{\sqrt{(1+2\chi/c^2)}}.$$

Since the spatial geometry may be regarded as Euclidean in a sufficiently small part of space, it follows from these equations that a light wave with sufficiently small wave-length, i.e. in the limit of geometrical optics, will propagate with the velocity

$$w = \frac{c}{n} = \frac{c}{\sqrt{(\epsilon\mu)}} = c\sqrt{\left(1+\frac{2\chi}{c^2}\right)} \tag{94}$$

in accordance with the equation (VIII. 70). The trajectory of a light ray in a static gravitational field is therefore determined in the same way as in an inhomogeneous refractive body, i.e. by Fermat's principle (see § 10), according to which the time which the light takes to travel between two points A and B in space is a minimum for the actual path chosen by the light ray. Mathematically this is expressed by the variational principle

$$\delta\int\limits_A^B \frac{d\sigma}{w} = \delta\int\limits_A^B \frac{d\sigma}{c\sqrt{(1+2\chi/c^2)}} = 0 \tag{95}$$

for all variations of the curve connecting A and B. Since $\gamma_{\iota\kappa} = g_{\iota\kappa}$ in our case, (95) may be written

$$\delta\int L(x^\iota, \dot{x}^\iota)\, d\lambda = 0 \tag{96}$$

with

$$L(x^\iota, \dot{x}^\iota) = \frac{\sqrt{(g_{\iota\kappa}\dot{x}^\iota\dot{x}^\kappa)}}{w(x^\iota)}, \qquad \dot{x}_\iota = \frac{dx^\iota}{d\lambda}. \tag{97}$$

Here, λ is an arbitrary parameter in a parametric representation $x^\iota = x^\iota(\lambda)$ of the curve considered. According to the equations (21), (24) in § 86, the condition (96) is equivalent to the Euler equations

$$\frac{d}{d\lambda}\left(\frac{g_{\iota\kappa}\dot{x}^\kappa}{w\sqrt{(g_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda)}}\right) = \frac{1}{2}\frac{(\partial g_{\kappa\lambda}/\partial x^\iota)\,\dot{x}^\kappa\dot{x}^\lambda}{w\sqrt{(g_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda)}} - \frac{\sqrt{(g_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda)}}{w^2}\frac{\partial w}{\partial x^\iota}. \tag{98}$$

Since the parameter λ was completely arbitrary, we may choose it so that

$$w\sqrt{(g_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda)} = \frac{d\sigma}{d\lambda}w = \text{constant} = 1 \tag{99}$$

along the curve determined by (98). These equations then reduce to

$$\frac{d_c u_\iota}{d\lambda} = \frac{d}{d\lambda}(g_{\iota\kappa}\dot{x}^\kappa) - \frac{1}{2}\frac{\partial g_{\kappa\lambda}}{\partial x^\iota}\dot{x}^\kappa\dot{x}^\lambda = -\frac{1}{w^3}\frac{\partial w}{\partial x^\iota} = -\frac{1}{w^4}\frac{\partial \chi}{\partial x^\iota}, \tag{100}$$

which together with (99) completely determine the trajectory of the light ray. The term on the right-hand side $-(1/w^4)\text{grad}\,\chi$ determines the deviation of the light ray from the 'straightest' line. According to (17) this term is equal to $(1/w^4)\mathbf{G}$, where $\mathbf{G}$ is the gravitational force on a unit mass, thus the light ray will be a straightest line only if the gravitational force $\mathbf{K}$ is zero.

The same result may be obtained from the equations (VIII. 87) for the time track of a light ray.† Since $\gamma_\iota = 0$, $w = c\sqrt{(-g_{44})}$, the last of these equations may be written

$$\left(\frac{d\sigma}{d\lambda}\right)^2 - w^2\left(\frac{dt}{d\lambda}\right)^2 = 0. \tag{101}$$

Further, we get from the first equations with $i = 4$ in the case of a static field

$$\frac{d}{d\lambda}\left(-\frac{w^2}{c^2}\frac{dx^4}{d\lambda}\right) = 0,$$

which shows that $w^2.\dot{t} =$ constant

along the time track. By a proper choice of λ this constant can be made equal to 1. Hence,

$$w^2.\dot{t} = 1, \tag{102}$$

and, on account of (101),

$$\left(\frac{d\sigma}{d\lambda}\right)^2 w^2 = 1,$$

which shows that the parameter λ defined by (102) is identical with the parameter defined by (99). For $i = 1, 2, 3$ we then get from (VIII. 87)

$$\frac{d}{d\lambda}(g_{\iota\kappa}\dot{x}^\kappa) = \frac{1}{2}\frac{\partial g_{\kappa\lambda}}{\partial x^\iota}\dot{x}^\kappa\dot{x}^\lambda - \frac{1}{2}\frac{\partial w^2}{\partial x^\iota}\dot{t}^2,$$

which, on account of (102), is identical with (100). Fermat's principle of least time thus holds in every static gravitational field.

† T. Levi-Civita, *Rend. Accad. Lincei* (5), No 26 (1917), p. 458; *Nuovo Cimento* (6), **16**, 105 (1918), H. Weyl, *Ann. d. Phys* **54**, 117 (1917).

XI

THE FUNDAMENTAL LAWS OF GRAVITATION IN THE GENERAL THEORY OF RELATIVITY

118. The gravitational field equations

IN the preceding sections we considered the influence of a given gravitational field on physical phenomena. We now turn to the most important problem in gravitational theory, which consists in finding the general equations determining the gravitational field variables $(\gamma_{\iota\kappa}, \gamma_\iota, \chi)$ or the g_{ik} from a given distribution of mass. After several attempts, this problem was finally solved by Einstein† in 1915. In Newton's theory of gravitation, the corresponding problem may be stated in the form of Poisson's equation

$$\Delta\chi = 4\pi k\overset{0}{\mu}, \tag{1}$$

where $$k = 6{\cdot}664\times 10^{-8}\ \text{cm.}^3\ \text{gm.}^{-1}\ \text{sec.}^{-2} \tag{2}$$

is the gravitational constant. This equation enables us to calculate the gravitational potential χ when the mass density $\overset{0}{\mu}$ is given as a function of the space coordinates.

On account of the equivalence of mass and energy, we must assume that any energy distribution, thus for instance an electromagnetic field, will create a gravitational field. Now, the energy density of any physical system is given by the component T_{44} of the energy-momentum tensor T_{ik} of the system, while $\chi = \frac{1}{2}c^2(-1-g_{44})$ is connected with the component g_{44} of the metric tensor; thus, equation (1) expresses the fact that a certain differential operator of second order, acting on g_{44}, is proportional to T_{44}. Since the general field equations must be covariant, and since different components of T_{ik} are mixed up by a transformation of coordinates, it is natural to assume that the general field equations are of the form

$$M_{ik} = -\kappa . T_{ik}, \tag{3}$$

where κ is a universal constant, and M_{ik} is a tensor of rank 2 depending on the metric tensor g_{ik} and its first and second derivatives only. Since the equation (3) for weak fields must reduce to Poisson's equation (1), M_{ik} must be linear in the second derivatives of g_{ik}, and the only possible expression for M_{ik} is then of the form

$$M_{ik} \equiv R_{ik} + c_1 R . g_{ik} + c_2 g_{ik}, \tag{4}$$

where c_1 and c_2 are constants, while R_{ik} and R are the contracted forms of the Riemann–Christoffel curvature tensor defined by (IX. 111, 114).

† A. Einstein, *Berl. Ber.*, pp. 778, 799, 844 (1915), *Ann. d. Phys.* **49**, 769 (1916).

On account of the symmetry properties of the tensors occurring in (4) the equations (3) represent ten differential equations for the ten functions g_{ik}. A simple consideration shows, however, that these ten equations cannot be independent (Hilbert).† Consider, for instance, the special case of empty space, where $T_{ik} = 0$. The equations (3) then reduce to

$$M_{ik} = 0. \tag{5}$$

If these equations were independent, the ten equations (5) would in a definite system of coordinates (x^i) allow us to determine uniquely the functions $g_{ik}(x^l)$ throughout the whole 4-space, when the values of g_{ik} and $\partial g_{ik}/\partial x^l$ are given on a hypersurface

$$x^4 = \text{constant} = a. \tag{6}$$

If we now introduce a new system of coordinates x'^i by

$$x'^i = x'^i(x), \qquad x^i = x^i(x'), \tag{7}$$

the transformed functions

$$g'_{ik} = \frac{\partial x^l}{\partial x'^i}\frac{\partial x^m}{\partial x'^k}g_{lm} \tag{8}$$

satisfy a set of differential equations

$$M'_{ik} = 0, \tag{9}$$

where, on account of the covariance of the equations (5), M'_{ik} are the same functions of g'_{ik}, $\partial g'_{ik}/\partial x'^l$, $\partial^2 g'_{ik}/\partial x'^l \partial x'^m$ as M_{ik} of g_{ik}, $\partial g_{ik}/\partial x^l$, $\partial^2 g_{ik}/\partial x^l \partial x^m$. Thus, by the same argument as before, the equations (9) enable us to determine uniquely the functions $g'_{ik}(x')$ from the values of g'_{ik} and $\partial g'_{ik}/\partial x'^l$ on the hypersurface $x^4 = x^4(x') = a$, and if we choose the transformations (7) such that $x'^i = x^i$ in the vicinity of the surface $x^4 = a$, but arbitrary elsewhere, we have

$$g'_{ik} = g_{ik}, \qquad \frac{\partial g'_{ik}}{\partial x'^l} = \frac{\partial g_{ik}}{\partial x^l}$$

on this surface, and $g'_{ik}(x'^l)$ must consequently be the same function of the variables (x'^l) as $g_{ik}(x^l)$ is of (x^l). This is, however, in contradiction to (8), which shows that the dependence of g'_{ik} on (x'^l) for points sufficiently far from the surface $x^4 = a$ will, in general, differ from the functional relationship between g_{ik} and (x^l).

Thus, if the field equations are to be covariant, we must assume that the quantities M_{ik} on the left-hand side of (5) and (3) satisfy four identities. This means that the solutions g_{ik} of the field equations contain four arbitrary functions corresponding to the four arbitrary functions in

† D. Hilbert, *Gött. Nachr.*, p. 395 (1915).

the transformations (7), which only change our space-time description, but not the physical system which produces the gravitational field. In fact, as seen in § 113, it is always possible by a proper choice of the space-time coordinates to ensure that the four functions g_{i4} have the prescribed values $-\delta_{i4}$ throughout space-time. The six independent equations which remain after the introduction of the four identities involving the quantities M_{ik} are thus just sufficient to determine the six remaining components $g_{\iota\kappa}$ of the metric tensor.

As we have seen in § 114, the theorems of conservation of energy and momentum of a material system in a general system of coordinates have the form

$$T^k_{i,k} = \operatorname{div}_i\{T^k_i\} = 0. \tag{10}$$

Thus, if we assume that the four identities involving the components of the tensor M_{ik} are

$$M^k_{i,k} = 0, \tag{11}$$

the conservation theorem for a material system is a consequence of the field equations (3) in the same way as the conservation of electric charge (X. 60) follows from Maxwell's equations (X. 58).

We therefore assume that the differential operators M_{ik} on the left-hand side of the field equations satisfy the four identities (11). According to (IX. 86) and (IX. 112) we now get from (4)

$$M^k_{i,k} = (\tfrac{1}{2}+c_1)(Rg^k_i)_{,k},$$

and, if this is to be identically zero, we must have

$$c_1 = -\tfrac{1}{2}.$$

Thus, putting $c_2 = -\lambda$, where λ is a universal constant, we get

$$M_{ik} = R_{ik}-\tfrac{1}{2}Rg_{ik}-\lambda g_{ik}, \tag{4'}$$

and the field equations (3) take the form

$$M_{ik} = R_{ik}-\tfrac{1}{2}Rg_{ik}-\lambda g_{ik} = -\kappa T_{ik}. \tag{12}$$

Writing down this tensor equation in mixed components, we get

$$M^k_i = R^k_i-\tfrac{1}{2}R\delta^k_i-\lambda\delta^k_i = -\kappa T^k_i. \tag{13}$$

By contraction we obtain, since $R^i_i = R$, $\delta^i_i = 4$,

$$R+4\lambda = \kappa T, \tag{14}$$

where

$$T = T^i_i \tag{15}$$

is the invariant obtained by contraction of the energy-momentum tensor T_{ik}. Eliminating R by means of (14) the field equations (12) can also be written in the form

$$R_{ik}+\lambda g_{ik} = -\kappa(T_{ik}-\tfrac{1}{2}Tg_{ik}). \tag{16}$$

119. The linear approximation for weak fields

The field equations (12) and (16) are in general *non-linear* partial differential equations in the functions g_{ik}. However, for weak gravitational fields, the field equations may approximately be replaced by a set of linear differential equations.† A weak field means that we can introduce a system of space-time coordinates in which the metric tensor is of the form

$$g_{ik} = G_{ik} + h_{ik}, \tag{17}$$

where G_{ik} is the constant metric tensor (VIII. 41) of the special theory, and the h_{ik} and their derivatives are small quantities whose squares may be neglected. The Christoffel symbols (IX. 50) will then obviously be small of the first order and, in the curvature tensor (IX. 99, 114), we can neglect all terms depending on the squares of the Christoffel symbols. Hence we get

$$\begin{aligned} R_{ik} &= \frac{\partial \Gamma^r_{ir}}{\partial x^k} - \frac{\partial \Gamma^r_{ik}}{\partial x^r} \\ &= \frac{G^{rs}}{2}\frac{\partial}{\partial x^k}\left(\frac{\partial h_{si}}{\partial x^r} + \frac{\partial h_{sr}}{\partial x^i} - \frac{\partial h_{ir}}{\partial x^s}\right) - \frac{G^{rs}}{2}\frac{\partial}{\partial x^r}\left(\frac{\partial h_{si}}{\partial x^k} + \frac{\partial h_{sk}}{\partial x^i} - \frac{\partial h_{ik}}{\partial x^s}\right) \\ &= \frac{G^{rs}}{2}\frac{\partial^2 h_{ik}}{\partial x^r \partial x^s} + \frac{1}{2}\left(\frac{\partial^2 h}{\partial x^i \partial x^k} - \frac{\partial^2 h^r_i}{\partial x^r \partial x^k} - \frac{\partial^2 h^r_k}{\partial x^i \partial x^r}\right), \end{aligned} \tag{18}$$

where we have put

$$h^r_i = G^{rs}h_{is}, \qquad h = h^r_r = G^{rs}h_{rs}. \tag{19}$$

Let us now first consider the case of a static distribution of matter. This means that among the systems of space-time coordinates of the type (17) there will be some in which the field variables h_{ik} and the components of the material energy-momentum tensor are independent of t. For the component R_{44} we then get from (18)

$$R_{44} = \frac{G^{\iota\kappa}}{2}\frac{\partial^2 h_{44}}{\partial x^\iota \partial x^\kappa} = -\frac{\Delta\chi}{c^2}, \tag{20}$$

where χ is the scalar gravitational potential defined by (VIII. 94) and Δ is Laplace's operator. If, in the case of matter at rest, we neglect the small contributions of the elastic stresses to the energy-momentum tensor we have, to a first approximation,

$$\left.\begin{aligned} T_{ik} &= G_{il}G_{km}\overset{0}{\mu}{}^0 U^l U^m = \delta_{i4}\delta_{k4}\overset{0}{\mu}{}^0 c^2 \\ T &= T^i_i = \overset{0}{\mu}{}^0 U_i U^i = -\overset{0}{\mu}{}^0 c^2 \\ T_{44} - \tfrac{1}{2}G_{44}T &= \tfrac{1}{2}\overset{0}{\mu}{}^0 c^2 \end{aligned}\right\}. \tag{21}$$

† A. Einstein, *Berl. Ber.*, p. 688 (1916).

Thus we get from (16) with $i = k = 4$

$$\Delta\chi - \lambda c^2 g_{44} = \frac{\kappa c^4}{2}\overset{0}{\mu}{}^0. \tag{22}$$

Now we know that Poisson's equation (1) is a good approximation for all static and quasi-static gravitational fields inside the solar system. We can therefore conclude that the constant λ must be so small that the λ-term may be neglected for all gravitational phenomena inside the planetary system of the sun. The λ-term can be of importance only for cosmological problems and in all other cases we shall therefore put λ equal to zero. The field equations (12) then reduce to

$$R_{ik} - \tfrac{1}{2}Rg_{ik} = -\kappa T_{ik}. \tag{23}$$

Further, comparison of (22) and (1) shows that the constant κ must be connected with the gravitational constant k by the equation

$$\kappa c^2 = \frac{8\pi k}{c^2} = 1{\cdot}87\times 10^{-27}\text{ cm gm.}^{-1} \tag{24}$$

Returning now to the case of a general weak field, we first remark that the system of coordinates corresponding to the form (17) with h_{ik} small to the first order is still to a large extent arbitrary. Any transformation of the type

$$x'^i = x^i + \delta x^i, \tag{25}$$

where $\delta x^i(x^l)$ is a function which is small to the first order, will lead to an expression for the transformed metric tensor of the same form as (17). Since the left-hand side of (23) is small to the first order, the same must be the case for κT_{ik}, which means that, to the approximation considered, T_{ik} may be treated as invariant under the transformations (25). Hence the components of T_{ik} may be put equal to the corresponding expressions in the local system of inertia.

The expression (18) for R_{ik} may now be written in the form

$$R_{ik} = \tfrac{1}{2}\Box h_{ik} - \frac{1}{2}\frac{\partial}{\partial x^i}\frac{\partial \chi_k^l}{\partial x^l} - \frac{1}{2}\frac{\partial}{\partial x^k}\frac{\partial \chi_i^l}{\partial x^l} \tag{26}$$

with $\chi_i^k = h_i^k - \tfrac{1}{2}\delta_i^k h, \qquad \chi_{ik} = h_{ik} - \tfrac{1}{2}G_{ik}h, \qquad h = h_l^l.$ (27)

The quantities χ_{ik}, χ_i^k transform like the components of a tensor by linear transformations of the space-time coordinates. On the other hand, it is always possible† by a transformation of the type (25) to ensure that the χ_i^k satisfy the equations

$$\frac{\partial \chi_i^k}{\partial x^k} = 0. \tag{28}$$

† See D. Hilbert, *Gött. Nachr.*, p. 53 (1917), and the solution (31) which satisfies (28).

For R_{ik} and R we then get

$$R_{ik} = \tfrac{1}{2}\Box h_{ik}, \qquad R = R_i^i = \tfrac{1}{2}\Box h, \tag{29}$$

and the field equations (23) reduce to

$$\Box \chi_{ik} = -2\kappa T_{ik}. \tag{30}$$

These equations are of the same form as the equations (V. 26) for the electromagnetic potentials, and the solutions of (30) which vanish at infinity are therefore analogous to the 'retarded' potentials (V. 46)

$$\chi_{ik}(x^\iota, t) = \frac{2\kappa}{4\pi}\int \frac{T_{ik}(x'^\iota, t-r/c)}{r}\,dx'^1dx'^2dx'^3, \tag{31}$$

where $r = \sqrt{\Big\{\sum_{\iota=1,2,3}(x^\iota-x'^\iota)^2\Big\}}$. If we can prove that the functions defined by (31) also satisfy the conditions (28), they must be the required solutions of the approximate field equations (23). The proof that the solutions (31) really satisfy the equations (28) runs exactly as the proof of the validity of the Lorentz condition for the retarded electromagnetic potentials. In the latter case, this proof is based on the law of conservation of electric charge (see V. 41). In the present case, the validity of (28) follows in the same way from the law of conservation of energy and momentum which, to the approximation considered, may be written

$$\frac{\partial T_i^k}{\partial x^k} = 0. \tag{32}$$

From (31) we can now also find explicit expressions for the quantities h_{ik}. First we get from (27)

$$\chi_l^l = h_l^l - 2h = -h, \qquad h_{ik} = \chi_{ik} + \tfrac{1}{2}G_{ik}h = \chi_{ik} - \tfrac{1}{2}G_{ik}\chi_l^l; \tag{33}$$

hence, by means of (31),

$$\chi_l^l = \frac{\kappa}{2\pi}\int \frac{T'\,dV'}{r}$$

and

$$h_{ik} = \frac{\kappa}{2\pi}\int \frac{T'_{ik}-\tfrac{1}{2}G_{ik}T'}{r}\,dV', \tag{34}$$

where the prime in T'_{ik} and T' means that these quantities have to be taken at the place (x'^ι) of the volume element $dV' = dx'^1dx'^2dx'^3$ and at the retarded time $t-r/c$.

120. Simple applications of the linear equations for weak fields. The relativity of centrifugal forces and Coriolis forces

To begin with let us consider a *static* distribution of matter, where the mass density $\mu^0 = \mu^0(x, y, z)$ is a given function of the space coordinates

$x^\iota = (x, y, z)$. In this case we have, according to (21) and (34),

$$\left.\begin{aligned} T_{\iota k} &= \delta_{\iota 4}\delta_{k4}\cdot\mu^0 c^2, \qquad T = -\mu^0 c^2 \\ h_{44} &= \frac{\kappa c^2}{4\pi}\int\frac{\mu^0(x',y',z')\,dx'dy'dz'}{|\mathbf{x}-\mathbf{x}'|} \\ h_{\iota 4} &= 0 \\ h_{\iota\kappa} &= \frac{\kappa c^2}{4\pi}\int\frac{\mu^0(x',y',z')\,dx'dy'dz'}{|\mathbf{x}-\mathbf{x}'|}\,\delta_{\iota\kappa} = h_{44}\,\delta_{\iota\kappa} \end{aligned}\right\}, \tag{35}$$

and for the gravitational potential we get, on account of (24), the usual expression of the Newtonian theory

$$\chi = -\frac{c^2}{2}h_{44} = -k\int\frac{\mu^0(\mathbf{x}')\,dV'}{|\mathbf{x}-\mathbf{x}'|}. \tag{36}$$

Hence the line element is of the form

$$\begin{aligned} ds^2 &= (G_{\iota k}+h_{\iota k})\,dx^\iota dx^k \\ &= \left(1-\frac{2\chi}{c^2}\right)(dx^2+dy^2+dz^2)-\left(1+\frac{2\chi}{c^2}\right)c^2\,dt^2, \end{aligned} \tag{37}$$

where χ is given by (36).

The spatial line element is

$$d\sigma^2 = \left(1-\frac{2\chi}{c^2}\right)(dx^2+dy^2+dz^2), \tag{38}$$

hence the geometry is only approximately Euclidean and the coordinates x, y, z are not exactly Cartesian. In general it is not possible by a change of the space coordinates to introduce Cartesian coordinates. However, the deviations from Euclidean geometry are in most cases too small to be measured. At the surface of the earth, for instance, the quantity $2\chi/c^2$ is of the order of magnitude 10^{-9}.

For a system of material particles with the masses $M_1^0, M_2^0,\ldots$ situated at the places $\mathbf{X}_1, \mathbf{X}_2,\ldots$, respectively, we get from (36)

$$\chi(\mathbf{x}) = -k\left(\frac{M_1^0}{|\mathbf{x}-\mathbf{X}_1|}+\frac{M_2^0}{|\mathbf{x}-\mathbf{X}_2|}+\ldots\right). \tag{39}$$

For a single particle the potential and the line element are spherically symmetrical. We have

$$\left.\begin{aligned} ds^2 &= \left(1+\frac{2kM^0}{c^2 r}\right)(dx^2+dy^2+dz^2)-\left(1-\frac{2kM^0}{c^2 r}\right)c^2\,dt^2 \\ r &= |\mathbf{x}-\mathbf{X}| \end{aligned}\right\}. \tag{40}$$

In the same way, the case of a *stationary* mass current distribution may be treated by means of Einstein's approximate equations (34). Thirring and Lense† calculated, for instance, the influence of the rotation of a central astronomical body on the gravitational field and the corresponding effects on the motion of the satellites. All such effects are too small to be observed‡ and we shall not consider them here.

There is one effect of this kind, however, which, although small, is of theoretical importance since it throws new light on the nature and origin of the centrifugal and Coriolis forces appearing in a rotating system of coordinates S. According to the idea of Einstein underlying the general principle of relativity (cf. § 82), these forces are gravitational forces originating from the rotation of the distant celestial masses relative to S, and such 'non-permanent' gravitational fields should satisfy the same general field equations as the permanent fields. The approximate solutions (34) for weak fields do not directly allow us to treat the effects of the distant celestial masses, but we may expect that a rotating spherical shell of uniform mass density will produce effects inside the shell similar to the rotation of the distant celestial masses.

If the shell is at rest, the potential χ given by (36) is constant inside the shell and has the value $\chi = -kM_0/R$, i.e.

$$\frac{2\chi}{c^2} = -\frac{\kappa c^2}{4\pi}\cdot\frac{M_0}{R}, \tag{41}$$

where M_0 is the total mass and R the radius of the shell.

Apart from the constants $1-2\chi/c^2$ and $1+2\chi/c^2$, which can be removed by a simple change of scale of the space and time coordinates, the line element (37) thus has the same form inside the shell as in the special theory of relativity. In the case of a shell moving with constant velocity, which can be reduced to the former case by a Lorentz transformation, the line element inside the shell must therefore again be of the special relativity type. This is also easily seen directly from (34), using the expression for T_{ik} in the case of stationary rectilinear motion of the matter. This illustrates the fact that the uniform rectilinear motion of the distant celestial masses relative to the different systems of inertia does not give rise to any gravitational forces in these systems.

For a rotating shell of matter, however, Thirring§ found the interesting result that the field in the interior of the shell, as determined by the equations (30), is similar to the field in a rotating system of coordinates,

† H. Thirring and J. Lense, *Phys. ZS* **19**, 156 (1918).
‡ De Sitter, *Monthly Notices*, **76**, 699 (1916), **77**, 155 (1916).
§ H. Thirring, *Phys. ZS.* **19**, 33 (1918); **22**, 29 (1921).

thus leading to gravitational forces similar to the usual centrifugal and Coriolis forces. We shall here consider the somewhat simpler case of a rotating massive ring of rest mass M_0 and radius R, which is rotating clockwise in the xy-plane with angular velocity ω. If $(x', y', 0)$ are the coordinates of a point on the ring, we have

$$\left.\begin{aligned} &R^2 = x'^2+y'^2 \\ &T_{\iota\kappa} = T^{\iota\kappa} = \frac{\mu^0 u^\iota u^\kappa}{1-R^2\omega^2/c^2}, \qquad T_{\iota 4} = T_{4\iota} = -T^{\iota 4} = -\frac{\mu^0 u^\iota c}{1-R^2\omega^2/c^2} \\ &T_{44} = T^{44} = \frac{\mu^0 c^2}{1-R^2\omega^2/c^2}, \qquad T = T^i_\iota = T_{\iota\iota}-T_{44} = -\mu^0 c^2 \\ &u^\iota = \omega(y', -x', 0) \end{aligned}\right\}. \tag{42}$$

From (IV. 198) we get for the total rest mass M_0 of the ring

$$M_0 = \int \mu_0\, dV = \frac{1}{\sqrt{(1-R^2\omega^2/c^2)}} \int \mu^0\, dV.$$

Hence, by means of (34),

$$h_{\iota k}(x,y,z) = C \int \frac{t'_{\iota k}\, ds}{\rho}, \tag{43}$$

where

$$\rho = \sqrt{\{(x-x')^2+(y-y')^2+z^2\}} \tag{44}$$

is the distance between the point (x,y,z) and the line element ds on the ring. Further,

$$C = \frac{M_0 \kappa c^2}{4\pi^2 R\sqrt{(1-R^2\omega^2/c^2)}} \tag{45}$$

and

$$t'_{\iota k} = \begin{bmatrix} \frac{\omega^2}{c^2}y'^2+\frac{1}{2}-\frac{1}{2}\frac{R^2\omega^2}{c^2} & -\frac{x'y'}{c^2}\omega^2 & 0 & -\frac{\omega y'}{c} \\ -\frac{y'x'}{c^2}\omega^2 & \frac{\omega^2}{c^2}x'^2+\frac{1}{2}-\frac{1}{2}\frac{R^2\omega^2}{c^2} & 0 & \frac{\omega x'}{c} \\ 0 & 0 & \frac{1}{2}-\frac{1}{2}\frac{R^2\omega^2}{c^2} & 0 \\ -\frac{\omega y'}{c} & \frac{\omega x'}{c} & 0 & \frac{1}{2}+\frac{1}{2}\frac{R^2\omega^2}{c^2} \end{bmatrix}. \tag{45'}$$

We shall confine ourselves to the consideration of points (x, y, z) whose distance from the origin is small compared with R. This means that we can use an expansion of $1/\rho$ in terms of x/R, y/R, z/R. Neglecting all terms of higher than second order in these quantities and putting

$$x' = R\cos\vartheta', \qquad y' = R\sin\vartheta',$$

we get

$$\frac{1}{\rho} = \frac{1}{R}\left(1 - \frac{1}{2}\frac{x^2+y^2+z^2}{R^2} + \frac{x}{R}\cos\vartheta' + \frac{y}{R}\sin\vartheta' + \frac{3}{2}\frac{x^2}{R^2}\cos^2\vartheta' + \right.$$
$$\left. + \frac{3}{2}\frac{y^2}{R^2}\sin^2\vartheta' + \frac{3}{2}\frac{xy}{R^2}\sin 2\vartheta'\right). \quad (46)$$

Using this together with $ds = R\,d\vartheta'$ in (43) the integration over ϑ' from 0 to 2π can be easily performed. We shall only write down the expressions obtained for the components $g_{\iota 4} = G_{\iota 4}+h_{\iota 4}$, which determine the dynamical potentials, i.e.

$$\left.\begin{aligned} g_{14} &= -\frac{C\pi\omega}{c}y = \frac{M\kappa c^2}{4\pi R}\left(-\frac{\omega y}{c}\right) \\ g_{24} &= \frac{M\kappa c^2}{4\pi R}\frac{\omega x}{c} \\ g_{34} &= 0 \\ g_{44} &= -1+\frac{M\kappa c^2}{4\pi R}\left(1+\frac{x^2+y^2-2z^2}{4R^2}+\frac{R^2\omega^2}{c^2}\right)+\frac{M\kappa c^2}{16\pi R}\frac{\omega^2}{c^2}(x^2+y^2-2z^2) \end{aligned}\right\}, \quad (47)$$

where $$M = M_0(1-R^2\omega^2/c^2)^{-\frac{1}{2}}$$

is the relativistic mass of the shell. These expressions should hold for all values of ω for which $R\omega/c < 1$, the only approximations introduced being those arising from the use of the equations for weak fields and the expansion of $1/\rho$ in terms of x/R, y/R, z/R.

For a ring of mass M at rest we get in the same way

$$g_{\iota 4} = \left[-1+\frac{M\kappa c^2}{4\pi R}\left(1+\frac{x^2+y^2-2z^2}{4R^2}\right)\right]\delta_{\iota 4}.$$

In this case, the scalar potential is thus not constant as inside a closed shell. Hence, the contribution to the dynamical potentials arising from the rotation of the ring is, apart from the unimportant constant term

$$\frac{M\kappa c^2}{4\pi R}\frac{R^2\omega^2}{c^2}$$

in the expression for g_{44},

$$\left.\begin{aligned} \gamma_\iota &= \frac{M\kappa c^2}{4\pi R}\frac{\omega}{c}(-y,x,0) \\ \chi &= -\frac{M\kappa c^2}{16\pi R}\frac{\omega^2}{2}(x^2+y^2-2z^2) \end{aligned}\right\}. \quad (48)$$

A comparison with (X. 18′) shows that the gravitational force on a moving test particle inside a heavy wheel rotating clockwise in the xy-plane is of the same type as in a system of coordinates rotating counter-clockwise. It is true that on account of the smallness of the quantity $M\kappa c^2/4\pi R$ the effect is too small to be measured, which explains the negative result of the experiment actually performed by Friedlander,† but anyhow the above considerations suggest a connexion between the gravitational constant κ, the total mass M in the world, and the mean distance R of the distant celestial masses, of the type

$$\frac{M\kappa c^2}{4\pi R} \approx 1. \tag{49}$$

It is interesting that the dependence on the angular velocity of the gravitational forces inside a rotating wheel is exactly the same as in a rotating system of reference. The vector potentials γ_ι in (48), which give rise to the Coriolis forces, are even of the usual form as regards their dependence on the coordinates (x, y, z). On the other hand, the scalar potential in (48) contains besides the usual term x^2+y^2 a term depending on z^2. This term gives rise to an axial component of the 'centrifugal' force which tends to drag a test particle into the plane of the rotating wheel. One could think that this unexpected deviation from the usual centrifugal force is due to the particular mass distribution which we have considered. However, the calculations of Thirring show that the same effect appears inside a homogeneous rotating shell. The purely radial character of the usual centrifugal force therefore rather indicates that the approximate field equations (30), which require explicit assumptions regarding the boundary conditions at infinity in order to give unique solutions, do not give an adequate description of the world as a whole. Firstly, the exact equations (12) are non-linear and, secondly, they contain the λ-term which for cosmological distances may be of importance. In fact, this term entirely changes the character of the field equations, in particular as regards the question of the boundary conditions (see § 132). But even if we confine ourselves to the linear approximation of weak fields, a simple calculation shows that the use of the λ-term introduces into the solutions of the equations for the case of a rotating wheel extra terms which are of the same type as those occurring in (48), but multiplied by $-\lambda R^2$. The terms depending on (x^2+y^2) and z^2 do not, however, have the same ratio as in (48); therefore, if

$$\lambda R^2 \approx 1, \tag{50}$$

† B. and T. Friedländer, *Absolute und relative Bewegung*, Berlin, 1896.

it is understandable that the terms in χ which contain z^2 may cancel and that we are left with a purely radial centrifugal force. This consideration is, of course, very rough, since the linear equations probably represent a bad approximation when applied to the world as a whole; but, as we shall see in § 133, the relations (49) and (50) also follow from Einstein's solution of the exact gravitational equations applied to an ideal model of the universe.

Finally, if we treat the case of a *non-stationary* distribution of matter by means of (30) and (34), the close analogy to the wave equations of electrodynamics leads at once to the result that fluctuating matter in general gives rise to the emission of gravitational waves travelling with the velocity of light and carrying with them a certain amount of energy. As shown by Einstein,† the gravitational energy (see § 126) emitted in this way is, however, too small to give any measurable astronomical effect.

121. Equivalent systems of coordinates. Systems with spherical symmetry

Let (x^i) be an arbitrary system of coordinates with the metric tensor $g_{ik} = g_{ik}(x^l)$. If we introduce a new system of coordinates x'^i by

$$x'^i = x'^i(x^l), \tag{51}$$

the transformed components of the metric tensor

$$g'_{ik}(x'^r) = \frac{\partial x^l}{\partial x'^i}\frac{\partial x^m}{\partial x'^k} g_{lm}(x^s) \tag{52}$$

are not in general form-invariant functions of the coordinates, i.e. g'_{ik} is generally not the same function of the transformed variables (x'^l) as are the g_{ik} of the variables (x^l).

Two systems of coordinates (x^i) and (x'^i) for which the components of the metric tensor are form-invariant functions of the space-time coordinates under the transformation (51) may be called *equivalent*, since any physical process will have the same course of development in the two systems. The existence of equivalent systems of coordinates imposes a certain condition on the gravitational field, since the functions $g_{ik}(x^l)$ obviously must satisfy the functional equations

$$g_{ik}(x'^r(x^s)) = \frac{\partial x^l}{\partial x'^i}\frac{\partial x^m}{\partial x'^k} g_{lm}(x^s). \tag{53}$$

In some cases the gravitational field variables g_{ik} are form-invariant under a whole group of transformations. This is, for instance, the case

† A. Einstein, *Berl. Ber*, p. 154 (1918).

for all non-permanent gravitational fields, for, if we first introduce pseudo-Cartesian coordinates X^ι by the transformations

$$X^\iota = f^\iota(x^l), \tag{54}$$

we have $$ds^2 = g_{\iota k}\, dx^\iota dx^k = G_{\iota k}\, dX^\iota dX^k. \tag{55}$$

Further, by performing a Lorentz transformation

$$X'^\iota = \mathrm{A}^\iota_k\, X^k \tag{56}$$

and afterwards introducing new space-time coordinates (x'^ι) by means of the transformations

$$X'^\iota = f^\iota(x'^l) \tag{57}$$

with the same functions f^ι as in (54), we have

$$g_{\iota k}\, dx^\iota dx^k = G_{\iota k}\, dX^\iota dX^k = G'_{\iota k}\, dX'^\iota dX'^k = g'_{\iota k}\, dx'^\iota dx'^k,$$

and, since $$G'_{\iota k} = G_{\iota k} = \text{constant},$$

$g'_{\iota k}$ must be the same functions of the (x'^l) as are the $g_{\iota k}$ of (x^l). The transformation $x'^\iota = x'^\iota(x^l)$ defined by (54), (56), and (57), which may be called a *generalized Lorentz transformation*, thus connects two equivalent systems of coordinates, and the components $g_{\iota k}$ of the metric tensor in a non-permanent gravitational field are form-invariant under the group of generalized Lorentz transformations.

In the case of permanent gravitational fields, it is generally not possible to introduce such space-time coordinates x^ι that the $g_{\iota k}$ are form-invariant under the group of four-dimensional orthogonal transformations, but in some important cases the gravitational potentials are form-invariant under the sub-group of spatial orthogonal transformations. Such systems are naturally called *spherically symmetric*. Putting $x^\iota = (\mathbf{x}, ct) = (x, y, z, ct)$, the $g_{\iota k}$ are then form-invariant under any orthogonal transformation of the three variables x, y, z with constant t. Generally the coordinates (x, y, z) will not be Cartesian, the spatial geometry being in general non-Euclidean, nevertheless, the line element $ds^2 = g_{ik}\, dx^\iota dx^k$ can in this case be a function only of the well-known form-invariants of the group of three-dimensional rotations in a Euclidean space. These invariants are

$$\left.\begin{aligned} r &= \{x^2+y^2+z^2\}^{\frac{1}{2}}, \quad dr \\ dx^2+dy^2+dz^2 &= dr^2+r^2\, d\theta^2+r^2\sin^2\theta\, d\phi^2 \\ x\,dx+y\,dy+z\,dz &= r\,dr, \quad dt, \quad \text{and} \quad t \end{aligned}\right\}. \tag{58}$$

Since the line element is a quadratic form in the differentials, the most general possible expression for ds^2 in a system with spherical symmetry is therefore

$$ds^2 = F(r,t)\,dr^2 + G(r,t)(r^2\,d\theta^2 + r^2\sin^2\theta\,d\phi^2) + \\ + 2H(r,t)\,drdt + L(r,t)\,dt^2, \tag{59}$$

where F, G, H, L are functions of r and t, only.

This expression may be further reduced by a suitable choice of coordinates. Introducing a new variable r' instead of r by the transformation

$$r'^2 = r^2 G(r,t), \tag{60}$$

we see that the line element in the new variables is of the form

$$ds^2 = M(r',t)\,dr'^2 + r'^2(d\theta^2 + \sin^2\theta\,d\phi^2) + 2N(r',t)\,dr'dt + O(r',t)\,dt^2. \tag{61}$$

The quantities $\omega_{\iota\kappa}$ defined by (VIII. 110) are obviously zero in the present case. Hence it is possible by a simple change of rate of the coordinate clocks to make the vector potentials disappear, and since $N(r',t)$ does not depend on θ and ϕ, this may be obtained by a time transformation of the form

$$t' = f(r',t), \tag{62}$$

which will not affect the second term in (61). Dropping the primes, the line element may thus be expressed in the standard form

$$ds^2 = a\,dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2) - bc^2\,dt^2, \tag{63}$$

where

$$a = a(r,t), \qquad b = b(r,t) = 1 + \frac{2\chi}{c^2} \tag{64}$$

are functions of r and t which, on account of (VIII. 52), must be positive for all values of r and t.

122. Static systems with spherical symmetry

If the system is *static* and *spherically symmetrical*, the functions a and b are independent of t, and in this case the components of the tensor M_{ik} defined by (4′) are easily calculated. With $x^i = (r, \theta, \phi, ct)$ we have

$$g_{11} = a(r), \qquad g_{22} = r^2, \qquad g_{33} = r^2\sin^2\theta, \qquad g_{44} = -b, \tag{65}$$

all other components vanishing. The coordinate system is orthogonal; the non-vanishing components of g^{ik} are therefore simply

$$\left.\begin{aligned} g^{11} &= \frac{1}{g_{11}} = \frac{1}{a}, & g^{22} &= \frac{1}{g_{22}} = \frac{1}{r^2} \\ g^{33} &= \frac{1}{g_{33}} = \frac{1}{r^2\sin^2\theta}, & g^{44} &= \frac{1}{g_{44}} = -\frac{1}{b} \end{aligned}\right\} \tag{66}$$

and the Christoffel symbols (IX. 50) are of the form

$$\Gamma^{i}_{kl} = \frac{1}{g_{ii}}\frac{1}{2}\left(\delta_{ik}\frac{\partial g_{ii}}{\partial x^{l}} + \delta_{il}\frac{\partial g_{ii}}{\partial x^{k}} - \delta_{kl}\frac{\partial g_{kk}}{\partial x^{i}}\right) \tag{67}$$

(no summation over i and k !).

Since the g_{ik} are independent of ϕ and x^4, we see that the only independent non-vanishing components of $\Gamma^{i}_{kl} = \Gamma^{i}_{lk}$ are the following:

$$\left.\begin{array}{l}\Gamma^{1}_{11} = \frac{a'}{2a},\quad \Gamma^{1}_{22} = -\frac{r}{a},\quad \Gamma^{1}_{33} = -\frac{r}{a}\sin^{2}\theta,\quad \Gamma^{1}_{44} = \frac{b'}{2a} \\ \Gamma^{2}_{12} = \frac{1}{r},\quad \Gamma^{2}_{33} = -\sin\theta\cos\theta,\quad \Gamma^{3}_{13} = \frac{1}{r},\quad \Gamma^{3}_{23} = \cot\theta,\quad \Gamma^{4}_{14} = \frac{b'}{2b}\end{array}\right\}, \tag{68}$$

where the accents denote differentiation with respect to r.

Using (68) in (IX. 114) we obtain the tensor R_{ik} from which we get the scalar $R = R^{i}_{i}$ by contraction. Finally, the tensor M_{ik} and the corresponding mixed components M^{k}_{i} are obtained from (4′). A straightforward consideration shows that the non-diagonal elements of this tensor are zero and the four diagonal elements M^{1}_{1}, M^{2}_{2}, M^{3}_{3}, M^{4}_{4} are functions of r only. From the transformation properties of tensor components under rotations it then follows that the components M^{2}_{2} and M^{3}_{3} must be equal. However, on account of the identities (11), even the three quantities M^{1}_{1}, $M^{2}_{2} = M^{3}_{3}$, M^{4}_{4} cannot be independent. As is easily seen, the equations (11) reduce for a static spherical system to essentially one equation, only, by means of which the components $M^{2}_{2} = M^{3}_{3}$ may be expressed in terms of M^{1}_{1}, M^{4}_{4}, and dM^{1}_{1}/dr. The calculations give the following expressions for the non-vanishing components of M^{k}_{i}.

$$M^{1}_{1} = -\frac{b'}{abr} + \frac{1}{r^{2}}\left(1 - \frac{1}{a}\right) - \lambda, \tag{69 a}$$

$$M^{4}_{4} = \frac{a'}{a^{2}r} + \frac{1}{r^{2}}\left(1 - \frac{1}{a}\right) - \lambda, \tag{69 b}$$

$$M^{2}_{2} = M^{3}_{3} = -\frac{1}{2a}\left[\left(\frac{b'}{b}\right)' - \frac{1}{2}\frac{a'}{a}\frac{b'}{b} + \frac{1}{2}\left(\frac{b'}{b}\right)^{2} + \frac{b'}{br} - \frac{a'}{ar}\right] - \lambda. \tag{69 c}$$

These explicit expressions for the components of the tensor M^{k}_{i} are easily seen to be in accordance with the identities (11).

123. Schwarzschild's exterior solution

In the empty space surrounding a material particle of mass M we have $T_i^k = 0$ and the field equations (13) reduce to

$$-\frac{b'}{abr}+\frac{1}{r^2}\left(1-\frac{1}{a}\right)-\lambda = 0, \tag{70a}$$

$$\frac{a'}{a^2r}+\frac{1}{r^2}\left(1-\frac{1}{a}\right)-\lambda = 0, \tag{70b}$$

$$-\frac{1}{2a}\left[\left(\frac{b'}{b}\right)'-\frac{1}{2}\frac{a'}{a}\frac{b'}{b}+\frac{1}{2}\left(\frac{b'}{b}\right)^2+\frac{b'}{br}-\frac{a'}{ar}\right]-\lambda = 0, \tag{70c}$$

the last equation being actually a consequence of the two first equations. From (70a) and (70b) we get by subtraction

$$\frac{a'b+ab'}{a^2br} = 0,$$

i.e.
$$(ab)' = 0 \tag{71}$$

or
$$ab = \text{constant}. \tag{72}$$

Putting
$$y = \frac{1}{a} \tag{73}$$

(70b) becomes, after multiplication by $-r^2$,

$$ry'+y-1+\lambda r^2 = 0 \tag{74}$$

or
$$(yr)' = 1-\lambda r^2. \tag{75}$$

By integration we thus get

$$ry = r-\frac{\lambda r^3}{3}-2m, \tag{76}$$

where m is a constant of integration, or

$$y = 1-\frac{2m}{r}-\frac{\lambda r^2}{3}. \tag{77}$$

From (73) and (77) we get the following solution of (70b).

$$a = \frac{1}{1-2m/r-\lambda r^2/3}. \tag{78}$$

The spatial line element which in our case is equal to the spatial part of (63) is thus

$$d\sigma^2 = \frac{dr^2}{1-2m/r-\lambda r^2/3}+r^2(d\theta^2+\sin^2\theta\, d\phi^2). \tag{79}$$

If we neglect the small λ-term, $d\sigma^2$ will in the limit of large r go over into the usual line element for a Euclidean space in polar coordinates, and it is interesting that the condition of spherical symmetry is sufficient to secure this result without any explicit use of boundary conditions at infinity. This result is, of course, also partly connected with the normalization (60) of the variable r which has been chosen so that the geometry on a surface $r =$ constant is the same as on a sphere of radius r in a Euclidean space. In the actual space (79), r will, however, not simply be the radial distance, the distance between two points (r_1, θ, ϕ) and (r_2, θ, ϕ) measured with standard measuring-rods being

$$l = \int\limits_{r_1}^{r_2} (1-2m/r-\lambda r^2/3)^{-\frac{1}{2}}\, dr. \tag{80}$$

An observer at great distance from the central body will, however, try to make a picture of the system in a Euclidean space; in this picture the quantity r plays the role of the distance from the centre, and the distance l measured by standard measuring-rods has no real importance in astronomy. The coordinates (r, θ, ϕ) may therefore be taken to be the usual polar coordinates applied in celestial mechanics.

According to (72) we have

$$b = \frac{\text{constant}}{a} = \text{constant} \times \left(1-\frac{2m}{r}-\frac{\lambda r^2}{3}\right). \tag{81}$$

By a simple change of the time-scale the constant may be made equal to 1, and we thus arrive at Schwarzschild's exterior solution†

$$ds^2 = \frac{dr^2}{1-2m/r-\lambda r^2/3}+r^2(d\theta^2+\sin^2\theta\, d\phi^2)-\left(1-\frac{2m}{r}-\frac{\lambda r^2}{3}\right)c^2\, dt^2, \tag{82}$$

or, if we neglect the λ-term which is of importance only for very large values of r,

$$ds^2 = \frac{dr^2}{1-2m/r}+r^2(d\theta^2+\sin^2\theta\, d\phi^2)-\left(1-\frac{2m}{r}\right)c^2\, dt^2. \tag{83}$$

This expression must hold outside a spherical distribution of matter and, since it is singular at a distance $r = r_0$ determined by the equation

$$1-\frac{2m}{r_0} = 0, \tag{84}$$

we may conclude that the 'radius' of the mass must be larger than the value determined by (84). The singularity at $r = r_0$ cannot be completely removed by the use of other 'static' coordinates. It can, however, be

† K. Schwarzschild, *Berl. Ber*, p. 189 (1916)

modified, for instance, by using 'isotropic' coordinates r', θ, ϕ, t defined by the transformation

$$r = r'\left(1+\frac{m}{2r'}\right)^2, \qquad r' = \tfrac{1}{2}\{(r^2-2mr)^{\frac{1}{2}}+r-m\}. \tag{85}$$

The line element then takes the form

$$ds^2 = \left(1+\frac{m}{2r'}\right)^4(dr'^2+r'^2\,d\theta^2+r'^2\sin^2\theta\,d\phi^2)-\frac{(1-m/2r')^2}{(1+m/2r')^2}c^2\,dt^2. \tag{86}$$

The singular point $r_0 = 2m$ has the new radial coordinate $r_0' = \frac{1}{2}m$, and it is seen that in (86) the singularity has been removed from the spatial part of the line element, but it appears again in g_{44} which vanishes at the same place. As shown by Serini, Einstein, and Pauli,† no non-singular solutions of the field equations for empty space exist which are stationary and have the form $g_{44} = -1+\text{constant}/r$ at infinity. As we shall see presently, this form of g_{44} indicates that the field is produced by a mass distribution around the origin. The scalar potential which, of course, is invariant under the transformation (85) has the following forms in the two systems of coordinates corresponding to (83) and (86)

$$\left.\begin{aligned} \chi &= (-g_{44}-1)\frac{c^2}{2} = -\frac{mc^2}{r}, \\ \chi' &= (-g'_{44}-1)\frac{c^2}{2} = -\frac{mc^2}{r'(1+m/2r')^2} \end{aligned}\right\}. \tag{87}$$

At large distances where the field is weak, both expressions reduce to the Newtonian form $-mc^2/r = -mc^2/r'$, which shows that the constant m must be connected with the mass M of the particle creating the field by the relation

$$m = \frac{kM}{c^2} = \frac{\kappa c^2 M}{8\pi}. \tag{88}$$

In the regions where the field is weak, (86) reduces to the expression

$$ds^2 = \left(1+\frac{2m}{r'}\right)(dx^2+dy^2+dz^2)-\left(1-\frac{2m}{r'}\right)c^2\,dt^2, \tag{86'}$$

where we have put

$$x = r'\sin\theta\cos\phi, \qquad y = r'\sin\theta\sin\phi, \qquad z = r'\cos\theta. \tag{86''}$$

On account of (88) this expression is identical with the static solution (40) of the linear approximate field equations (30)

As first remarked by Lemaître,‡ the singularity disappears in the line

† R Serini, *Atti Accad Lincei* (5), **27**[1], 235 (1918), A Einstein, *Revista* (*Univ Nac. Tucuman*), A, **2**, 11 (1941), A Einstein and W Pauli, *Ann. of Math.* **44**, 131 (1943).

‡ G E Lemaître, *Ann Soc. Scient. Bruxelles*, Sér A, **53**, 51 (1933), see also J L. Synge, *Proc Roy Irish Soc* **53**, No 6, 83 (1950)

element if we introduce a *non-static* system of coordinates r', θ, ϕ, t' by the transformations

$$r = (9m/2)^{\frac{1}{3}}(r'-ct')^{\frac{2}{3}}, \qquad dt' = dt - \frac{(2m/r)^{\frac{1}{2}}}{1-2m/r}\,dr. \tag{89}$$

The line element (83) then takes the form

$$ds^2 = \frac{2m}{r}dr'^2 + r^2(d\theta^2+\sin^2\theta\, d\phi^2) - c^2\,dt'^2, \tag{90}$$

where r depends on r' and t' by (89). The new system of coordinates is of the type considered in § 113, where the dynamical action of the gravitational field has been transformed away, the dynamical potentials (γ_ι, χ) being zero. The motion of a planet is described in these coordinates by the equations (X. 20′), i.e. the covariant components of the momentum vector of the particle at the time $t'+dt'$ are obtained from the corresponding vector at the time t' by parallel displacement in the space with the line element

$$d\sigma^2 = \frac{2m}{r}\,dr'^2 + r^2(d\theta^2+\sin^2\theta\, d\phi^2).$$

On account of the time-dependence of r the system of reference is not rigid, however, and the motion of the planet in these coordinates is therefore rather complicated.

124. Schwarzschild's solution for the interior of a perfect fluid

The energy-momentum tensor of a perfect fluid is given by (X. 48)

$$T_\iota^k = \left(\mathring{\mu}^0 + \frac{\mathring{p}}{c^2}\right)U_\iota\, U^k + \mathring{p}\,\delta_\iota^k. \tag{91}$$

On account of the field equations (13),

$$M_\iota^k = -\kappa T_\iota^k, \tag{92}$$

and the expressions (69) for the tensor M_ι^k, we see that a static spherically symmetric field of the type (63) is possible only if the velocity u^ι of the matter is zero and if the proper mass density $\mathring{\mu}^0$ and the proper pressure $\mathring{p}$ are functions of r only. Hence we have, according to (X. 2, 3) and (65),

$$\left.\begin{aligned} U^\iota &= \left(0,0,0,\frac{c}{\sqrt{b}}\right) \\ U_\iota &= g_{\iota k}\,U^k = (0,0,0,-c\sqrt{b}) \\ T_\iota^k &= -\left(\mathring{\mu}^0+\frac{\mathring{p}}{c^2}\right)c^2\,\delta_{\iota 4}\,\delta_{k4} + \mathring{p}\,\delta_\iota^k \\ \mathring{\mu}^0 &= \mathring{\mu}^0(r), \qquad \mathring{p} = \mathring{p}(r) \end{aligned}\right\}. \tag{92'}$$

Using this with the conservation law

$$T^k_{i,k} = \frac{1}{\sqrt{|g|}} \frac{\partial(\sqrt{|g|}\, T^k_i)}{\partial x^k} - \Gamma^r_{is} T^s_r = 0, \tag{93}$$

we get for $i = 1$

$$\frac{1}{\sqrt{|g|}} \frac{\partial}{\partial r}(\sqrt{|g|}\, \mathring{p}) + \Gamma^4_{14}(\mathring{\mu}^0 c^2 + \mathring{p}^0) - \Gamma^r_{1r}\, \mathring{p} = 0$$

or, by means of (68) and (IX. 69),

$$\frac{d\mathring{p}}{dr} + (\mathring{\mu}^0 c^2 + \mathring{p}) \frac{b'}{2b} = \frac{d\mathring{p}}{dr} + \frac{\mathring{\mu}^0 + \mathring{p}/c^2}{1+2\chi/c^2} \frac{d\chi}{dr} = 0. \tag{94}$$

This equation gives the dependence of the pressure on the scalar gravitational potential in the equilibrium state of a fluid under the influence of its own gravitational field. The other three equations of the conservation laws (93) do not give anything new.

The field equations (92) again reduce to only two independent equations for which we may take the equations

$$M^1_1 = -\kappa T^1_1, \qquad M^4_4 = -\kappa T^4_4, \tag{95}$$

the other equations (92) being consequences of (95) on account of the conservation equation (94). From (69 a, b), (92′), and (95) we thus get

$$\frac{b'}{abr} - \frac{1}{r^2}\left(1 - \frac{1}{a}\right) + \lambda = \kappa \mathring{p}, \tag{96a}$$

$$\frac{a'}{a^2 r} + \frac{1}{r^2}\left(1 - \frac{1}{a}\right) - \lambda = \kappa \mathring{\mu}^0 c^2. \tag{96b}$$

The equations (94), (96) together with the equation of state of the matter which gives the connexion between $\mathring{p}$ and $\mathring{\mu}^0$ determine the interior state and the gravitational field of the fluid.

For simplicity we shall assume that the fluid is practically incompressible. The proper mass density may then be treated as a constant and the solution of (96 b) may be obtained from the solution (78) of (70 b) by the substitution $\lambda \to \lambda + \kappa\mu^0 c^2$. Since our solution is to be regular as $r \to 0$, the constant of integration $2m$ in (78) must be put equal to zero. Hence we get

$$a = \frac{1}{1 - \frac{\lambda + \kappa\mathring{\mu}^0 c^2}{3} r^2} = \frac{1}{1 - \frac{r^2}{R^2}}, \tag{97}$$

where
$$R^2 = \frac{3}{\lambda + \kappa\mathring{\mu}^0 c^2}. \tag{98}$$

Further, on account of the constancy of $\overset{0}{\mu}{}^0$, we get at once by integration of (94)

$$(\overset{0}{\mu}{}^0 c^2+\overset{0}{p})\sqrt{b} = \text{constant}.$$

By addition of the equations (96 a, 96 b) and multiplication by $\sqrt{b}$, we therefore have

$$\frac{b'}{a\sqrt{b}\,r}+\frac{a'\sqrt{b}}{a^2 r} = \text{constant}.$$

Substituting in this equation the expression (97) for a we obtain

$$\sqrt{b}+R^2\frac{1-r^2/R^2}{r}\frac{d\sqrt{b}}{dr} = A, \tag{99}$$

where A is a constant. Putting $y = \sqrt{b}$ and introducing a new variable $x = (1-r^2/R^2)^{\frac{1}{2}}$ instead of r, (99) may also be written

$$y-x\frac{dy}{dx} = A. \tag{100}$$

The solution of (100) is $\qquad y = A-Bx,$

where B is a constant of integration.

Hence $$b = y^2 = (A-B\sqrt{(1-r^2/R^2)})^2. \tag{101}$$

Finally, using (97) and (101) with (96 a), we get the following expression for the pressure $\overset{0}{p}$ measured in a local system of inertia

$$\kappa\overset{0}{p} = \frac{3B\sqrt{(1-r^2/R^2)}-A}{R^2\{A-B\sqrt{(1-r^2/R^2)}\}}+\lambda. \tag{102}$$

Thus we arrive at Schwarzschild's interior solution†

$$ds^2 = \frac{dr^2}{1-r^2/R^2}+r^2(d\theta^2+\sin^2\theta\, d\phi^2)-\{A-B\sqrt{(1-r^2/R^2)}\}^2c^2\, dt^2. \tag{103}$$

The spatial geometry is defined by the line element

$$d\sigma^2 = \frac{dr^2}{1-r^2/R^2}+r^2(d\theta^2+\sin^2\theta\, d\phi^2). \tag{104}$$

Hence, the geometry on the surface $r = r_1 =$ constant is the same as on a sphere of radius r_1 in Euclidean space, but r_1 is not the distance to the origin $r = 0$ measured by standard rods, this distance being

$$l_1 = \int\limits_0^{r_1}\frac{dr}{\sqrt{(1-r^2/R^2)}} = R\sin^{-1}\frac{r_1}{R} = r_1\left(1+\frac{1}{6}\left(\frac{r_1}{R}\right)^2+\frac{3}{40}\left(\frac{r_1}{R}\right)^4+\dots\right). \tag{105}$$

† K. Schwarzschild, *Berl Ber*, p 424 (1916).

The volume of this sphere is

$$V_1 = \int_0^{r_1}\int_0^{\pi}\int_0^{2\pi} \sqrt{\gamma}\, dr d\theta d\phi = \int\int\int \frac{r^2 \sin\theta}{\sqrt{(1-r^2/R^2)}}\, dr d\theta d\phi = 4\pi \int_0^{r_1} \frac{r^2\, dr}{\sqrt{(1-r^2/R^2)}},$$

$$\text{i.e.}\quad V_1 = \frac{4\pi R^3}{2}\left[\sin^{-1}\frac{r_1}{R} - \frac{r_1}{R}\sqrt{\left(1-\frac{r_1^2}{R_1^2}\right)}\right] = \frac{4\pi r_1^3}{3}\left[1+\frac{3}{10}\left(\frac{r_1}{R}\right)^2+\ldots\right]. \quad (106)$$

Consider a sphere of fluid filling the space inside $r = r_1$ with a constant proper density of mass $\mathring{\mu}^0$. For $r < r_1$ we then have the solution (103), while Schwarzschild's exterior solution (82) must be valid for $r > r_1$. We now have to adjust the constants A and B so that (103) and (82) coincide for $r = r_1$, further, $\mathring{p}$ has to be zero at the surface of the sphere. If we neglect the λ-term, which anyhow has only a small effect inside the solar system, these conditions lead to the equations

$$1-\frac{2m}{r_1} = 1-\frac{r_1^2}{R^2} = \{A-B\sqrt{(1-r_1^2/R^2)}\}^2,$$

$$3B\sqrt{(1-r_1^2/R^2)}-A = 0, \qquad R^2 = \frac{3}{\kappa\mathring{\mu}^0 c^2}$$

with the solutions

$$\left.\begin{aligned} A &= \tfrac{3}{2}\sqrt{(1-r_1^2/R^2)}, \qquad B = \tfrac{1}{2}, \qquad R^2 = \frac{3}{\kappa\mathring{\mu}^0 c^2} \\ m &= \frac{r_1^3}{2R^2} = \frac{\kappa c^2\mathring{\mu}^0}{6} r_1^3 = \frac{k\mathring{\mu}^0}{c^2}\cdot\frac{4\pi}{3} r_1^3 \end{aligned}\right\}. \quad (107)$$

A comparison with (88) shows that the gravitational field of the spherical fluid at great distances corresponds to a mass

$$M = \frac{4\pi}{3} r_1^3 \mathring{\mu}^0, \quad (108)$$

which is thus the same as for a constant distribution of Newtonian mass over a sphere of radius r_1 in a Euclidean space.

According to (106) the real volume of the sphere is larger than $\frac{4\pi}{3}r_1^3$; on the other hand, $\mathring{\mu}^0$ was the mass density measured in a local system of inertia which differs from the mass density in the system of coordinates used here. Actually we shall show in § 128 that the quantity M is exactly equal to the total energy of the system divided by c^2. Anyhow, the difference between the real volume V_1 given by (106) and $\frac{4\pi}{3}r_1^3$ is in all astronomical applications very small. In the case of the sun, for instance, we may put

$$\mathring{\mu}^0 = 1{\cdot}4 \text{ gm. cm.}^{-3}, \qquad r_1 = 6{\cdot}95\times 10^{10} \text{ cm.} \quad (109)$$

Hence we get

$$\left.\begin{aligned} R &= 3{\cdot}5\times 10^{13}\text{ cm.} \\ \frac{r_1}{R} &\approx 2\times 10^{-3} \\ \frac{V_1-4\pi r_1^3/3}{4\pi r_1^3/3} &\approx 10^{-6} \end{aligned}\right\}. \tag{110}$$

Also the difference between the distance l_1 defined by (105) and the radial coordinate r_1 is far too small to be detected by the astronomical determination of r_1.

We also see that the condition $r_1 > r_0 = 2m$ for the applicability of the exterior solution (83) is amply satisfied; for from (107) we get

$$\frac{2m}{r_1} = \frac{r_1^2}{R^2} \approx 10^{-6} < 1.$$

As in the case of empty space we can also here introduce isotropic coordinates. For arbitrary functions $a(r)$, $b(r)$ in (63), this may be obtained by a transformation $r' = r'(r)$ satisfying the differential equation

$$\frac{dr'}{r'} = \sqrt{\{a(r)\}}\,\frac{dr}{r}, \tag{111}$$

the general solution of which is

$$r' = C\exp\left(\int^{r} \frac{\sqrt{\{a(r)\}}\,dr}{r}\right), \tag{112}$$

where C is an arbitrary constant. The line element then takes the form

$$\left.\begin{aligned} ds^2 &= \frac{r^2}{r'^2}[dr'^2+r'^2(d\theta^2+\sin^2\theta\, d\phi^2)]-bc^2\,dt^2 \\ &= \frac{r^2}{r'^2}(dx^2+dy^2+dz^2)-bc^2\,dt^2 \end{aligned}\right\}, \tag{113}$$

where (x, y, z) are given by (86″).

For $a(r)$ of the form (97) we thus get

$$r' = C\frac{r/R}{1+\sqrt{(1-r^2/R^2)}}, \qquad r = r'\frac{2CR}{C^2+r'^2} \tag{114}$$

and the line element inside an incompressible fluid takes the form

$$ds^2 = \frac{4C^2R^2}{(C^2+r'^2)^2}(dx^2+dy^2+dz^2)-\left[A-B\frac{C^2-r'^2}{C^2+r'^2}\right]^2 c^2\,dt^2. \tag{115}$$

The constant C can be determined so as to make (114) coincide with (85) at the boundary of the fluid.

Schwarzschild's solution represents the only exact solution of the gravitational field equations which has found any application in astronomy. Reissner† and Weyl‡ also solved the problem of the gravitational field produced by the electromagnetic energy in the surroundings of a charged particle. The result obtained by these authors is a line element of the form

$$ds^2 = \frac{dr^2}{1-2m/r+\kappa e^2/r^2}+r^2(d\theta^2+\sin^2\theta\, d\phi^2)-\left(1-\frac{2m}{r}+\frac{\kappa e^2}{r^2}\right)c^2\, dt^2. \tag{116}$$

The ratio of the two terms depending on the charge and the mass, respectively, is thus, on account of (88),

$$\frac{\kappa e^2}{r\,.\,2m} = \frac{4\pi e^2}{rMc^2}. \tag{117}$$

In the case of an electron, the two terms will therefore become of the same order of magnitude at a distance corresponding to the classical electron radius $a = e^2/Mc^2$. Both terms will, however, have negligible effects on the interaction between electrons as compared with the Coulomb interaction.

Further, exact solutions of the field equations for the case of arbitrary cylindrically symmetrical distributions of matter were given by Weyl§ and by Levi-Civita.‖

125. The variational principle for gravitational fields

Let Σ be an arbitrary domain in 4-space. Consider the four-dimensional invariant integral

$$J_1 = \int\limits_\Sigma R\, d\Sigma = \int\limits_\Sigma R_{ik}\, g^{ik}\sqrt{(-g)}\, dx^1dx^2dx^3dx^4, \tag{118}$$

where $R = R_{ik}\, g^{ik}$ is the curvature scalar and

$$R_{ik} = \frac{\partial\Gamma^l_{il}}{\partial x^k}-\frac{\partial\Gamma^l_{ik}}{\partial x^l}+\Gamma^r_{il}\Gamma^l_{kr}-\Gamma^r_{ik}\Gamma^l_{rl} \tag{119}$$

is the contracted curvature tensor. The integrand in (118) is an algebraic function of the g_{ik} and their derivatives. Since the contravariant components g^{ik} are uniquely determined by the g_{ik}, the integrand can also be expressed as a function of the g^{ik} and their derivatives. Let us now consider a variation of the metric tensor δg^{ik} which is arbitrary inside Σ,

† H. Reissner, *Ann. d. Phys.* **50**, 106 (1916).
‡ H. Weyl, ibid. **54**, 117 (1917); *Raum-Zeit-Materie*, 3rd ed., p. 223, Berlin, 1920.
§ H. Weyl, *Ann. d. Phys.* **54**, 117 (1917); **59**, 185 (1919).
‖ T. Levi-Civita, *Rend. Accad. Lincei* (5), No. 26 (1917), No. 27 (1918); No. 28 (1919).

but which vanishes together with the variation of the first derivatives at the boundary of Σ. The corresponding variation of J_1 is then

$$\delta J_1 = \int \delta R_{ik}\sqrt{(-g)}g^{ik}\,dx + \int R_{ik}\,\delta(\sqrt{(-g)}g^{ik})\,dx. \tag{120}$$

By variation of (119) we get

$$\delta R_{ik} = \frac{\partial\delta\Gamma^l_{il}}{\partial x^k} - \frac{\partial\delta\Gamma^l_{ik}}{\partial x^l} + \delta\Gamma^r_{il}.\Gamma^l_{kr} + \Gamma^r_{il}\,\delta\Gamma^l_{kr} - \delta\Gamma^r_{ik}.\Gamma^l_{rl} - \Gamma^r_{ik}\,\delta\Gamma^l_{rl}. \tag{121}$$

While the Christoffel symbols Γ^i_{kl} transform according to the transformation equations (IX. 53) which, on account of the first term on the right-hand side, differ from the transformation law of tensors, the variation $\delta\Gamma^i_{kl}$, which is the difference between two Christoffel symbols at the same point, will obviously transform like a tensor. Thus, the covariant derivative

$$(\delta\Gamma^i_{kl})_{,m} = \frac{\partial\delta\Gamma^i_{kl}}{\partial x^m} + \Gamma^i_{mr}\,\delta\Gamma^r_{kl} - \Gamma^r_{km}\,\delta\Gamma^i_{rl} - \Gamma^r_{lm}\,\delta\Gamma^i_{kr} \tag{122}$$

will also be a tensor. This allows us to write the tensor δR_{ik} in the simple form

$$\delta R_{ik} = (\delta\Gamma^l_{il})_{,k} - (\delta\Gamma^l_{ik})_{,l}. \tag{123}$$

This tensor relation is most easily proved by introducing a geodesic system of coordinates in which $\overset{0}{\Gamma}{}^i_{kl} = 0$; for in this system the right-hand sides of (121) and (123) are at once seen to be identical.

If we multiply (123) by $\sqrt{(-g)}g^{ik}$ we get, using (IX. 85, 86, 90),

$$\begin{aligned}\sqrt{(-g)}g^{ik}\,\delta R_{ik} &= \sqrt{(-g)}[(g^{ik}\,\delta\Gamma^l_{il})_{,k} - (g^{ik}\,\delta\Gamma^l_{ik})_{,l}] \\ &= \sqrt{(-g)}(g^{il}\,\delta\Gamma^k_{ik} - g^{ik}\,\delta\Gamma^l_{ik})_{,l} \\ &= \frac{\partial}{\partial x^l}(\sqrt{(-g)}[g^{il}\,\delta\Gamma^k_{ik} - g^{ik}\,\delta\Gamma^l_{ik}]).\end{aligned}$$

Hence the integrand of the first integral in (120) has the form of an ordinary divergence. This integral can therefore be transformed to an integral over the surface of Σ and, as the variations δg^{ik} and their first derivatives are zero on this surface, the first integral in (120) is zero.

Further, we get from (IX. 69′)

$$\delta\sqrt{(-g)} = \frac{1}{2}\frac{-\delta g}{\sqrt{(-g)}} = -\frac{1}{2}\frac{gg^{ik}}{\sqrt{(-g)}}\delta g_{ik} = -\tfrac{1}{2}\sqrt{(-g)}g_{ik}\,\delta g^{ik} \tag{124}$$

$$\delta\{\sqrt{(-g)}g^{ik}\} = \sqrt{(-g)}(\delta g^{ik} - \tfrac{1}{2}g^{ik}g_{lm}\,\delta g^{lm}) \tag{124'}$$

and the integrand of the second integral in (120) becomes

$$R_{ik}\,\delta\{\sqrt{(-g)}g^{ik}\} = \sqrt{(-g)}(R_{ik} - \tfrac{1}{2}Rg_{ik})\,\delta g^{ik}.$$

Thus we get for the variation of the invariant J_1

$$\delta J_1 = \int (R_{ik} - \tfrac{1}{2}Rg_{ik})\, \delta g^{ik}\sqrt{(-g)}\, dx. \tag{125}$$

Similarly, we obtain by means of (124)

$$\delta \int 2\lambda\sqrt{(-g)}\, dx = -\int \lambda g_{ik}\, \delta g^{ik}\sqrt{(-g)}\, dx. \tag{126}$$

Hence, adding (125) and (126), we see that the variation of the invariant

$$J = \int (R+2\lambda)\sqrt{(-g)}\, dx \tag{127}$$

is equal to

$$\delta J = \int M_{ik}\, \delta g^{ik}\sqrt{(-g)}\, dx, \tag{128}$$

where M_{ik} is the tensor appearing on the left-hand side of the field equations (12). The field equations in empty space

$$M_{ik} = 0 \tag{129}$$

are therefore equivalent to the variational principle

$$\delta J = 0 \tag{130}$$

for all variations where δg^{ik} and their first derivatives vanish at the boundary of Σ.

It is clear that (125) and (128) remain true if from J_1 and J we subtract any integral whose integrand has the form of an ordinary divergence, for this can be transformed into a surface integral which will give no contribution to δJ_1 and δJ for the variations in question. Now the contribution to $R\sqrt{(-g)} = R_{ik}\, g^{ik}\sqrt{(-g)}$ arising from the first two terms in (119) can be written

$$\left(\frac{\partial \Gamma^l_{il}}{\partial x^k} - \frac{\partial \Gamma^l_{ik}}{\partial x^l}\right)\sqrt{(-g)}g^{ik} = \frac{\partial}{\partial x^k}\{\sqrt{(-g)}g^{ik}\Gamma^l_{il}\} - \frac{\partial}{\partial x^l}(\{\sqrt{(-g)}g^{ik}\Gamma^l_{ik}\} -$$

$$-\frac{\partial\{\sqrt{(-g)}g^{ik}\}}{\partial x^k}\Gamma^l_{il} + \frac{\partial\{\sqrt{(-g)}g^{ik}\}}{\partial x^l}\Gamma^l_{ik}, \tag{131}$$

and, since the first two terms are ordinary divergences, they may thus be neglected. Further, substituting for $\partial g^{ik}/\partial x^l$ the expression (IX. 68) in terms of the Christoffel symbols, and using the relation

$$\frac{\partial\sqrt{(-g)}}{\partial x^l} = \sqrt{(-g)}\Gamma^r_{lr},$$

the last two terms in (131) are easily seen to reduce to $2\mathfrak{L}$ with

$$\mathfrak{L} \equiv \sqrt{(-g)}g^{ik}(\Gamma^r_{ik}\Gamma^l_{rl} - \Gamma^r_{il}\Gamma^l_{kr}). \tag{132}$$

Since the contribution to $R\sqrt{(-g)}$ from the last two terms in (119) is $-\mathfrak{L}$, we get from (125) and (128)

$$\delta\int \mathfrak{L}\,dx = \int \sqrt{(-g)}(R_{ik}-\tfrac{1}{2}Rg_{ik})\,\delta g^{ik}\,dx \tag{133}$$

and

$$\delta\int \mathfrak{H}\,dx = \int \sqrt{(-g)}M_{ik}\,\delta g^{ik}\,dx, \tag{134}$$

where

$$\mathfrak{H} \equiv \mathfrak{L}+2\lambda\sqrt{(-g)}. \tag{135}$$

The integrals $\int \mathfrak{L}\,dx$ and $\int \mathfrak{H}\,dx$ are not invariants. Nevertheless, since they are defined by the same equations (132), (135) in every system of coordinates, the relations (133) and (134) have a covariant meaning.

In contrast to the integrands of J_1 and J, $\mathfrak{L}$ and $\mathfrak{H}$ do not contain the second derivatives of the metrical tensor. They can therefore be regarded as functions of the quantities g^{ik} and their first derivatives

$$g_l^{ik} \equiv \frac{\partial g^{ik}}{\partial x^l}. \tag{136}$$

As $\delta g_l^{ik} = \dfrac{\partial}{\partial x^l}\delta g^{ik}$, we obviously have

$$\begin{aligned}\delta\int \mathfrak{L}\,dx &= \int \left(\frac{\partial\mathfrak{L}}{\partial g^{ik}}\delta g^{ik}+\frac{\partial\mathfrak{L}}{\partial g_l^{ik}}\delta g_l^{ik}\right)dx\\ &= \int \left[\frac{\partial\mathfrak{L}}{\partial g^{ik}}-\frac{\partial}{\partial x^l}\left(\frac{\partial\mathfrak{L}}{\partial g_l^{ik}}\right)\right]\delta g^{ik}\,dx\end{aligned} \tag{137}$$

and similarly

$$\delta\int \mathfrak{H}\,dx = \int \left[\frac{\partial\mathfrak{H}}{\partial g^{ik}}-\frac{\partial}{\partial x^l}\left(\frac{\partial\mathfrak{H}}{\partial g_l^{ik}}\right)\right]\delta g^{ik}\,dx. \tag{138}$$

Since the expressions (133) and (137) for $\delta\int \mathfrak{L}\,dx$ must be equal for any variation δg^{ik} inside the arbitrary domain Σ, we must have at every point

$$\sqrt{(-g)}(R_{ik}-\tfrac{1}{2}Rg_{ik}) = \frac{\partial\mathfrak{L}}{\partial g^{ik}}-\frac{\partial}{\partial x^l}\left(\frac{\partial\mathfrak{L}}{\partial g_l^{ik}}\right). \tag{139}$$

In the same way we get by comparison of (134) and (138)

$$\sqrt{(-g)}M_{ik} = \frac{\partial\mathfrak{H}}{\partial g^{ik}}-\frac{\partial}{\partial x^l}\left(\frac{\partial\mathfrak{H}}{\partial g_l^{ik}}\right). \tag{140}$$

In applying this argument we have treated the variables g^{ik}, g_l^{ik} as independent, thus disregarding the symmetry relations $g^{ik} = g^{ki}$, $g_l^{ik} = g_l^{ki}$. This is obviously permissible, provided that the result of the derivations on the right-hand sides of (139) and (140) is symmetrical in i and k. Although $\mathfrak{L}$ and $\mathfrak{H}$ in general have rather complicated transformation properties, the covariant expressions on the right-hand sides of (139) and (140) must transform like tensor densities.

If for constant g_l^{ik} all the variables g^{ik} are multiplied by a factor λ, the quantities g_{ik}, $\sqrt{(-g)}$, and Γ_{kl}^i are from (IX. 7, 4, 50) seen to be multiplied by the factors λ^{-1}, λ^{-2}, λ^{-1}, respectively. Hence, the quantity $\mathfrak{L}$ will be multiplied by the factor λ^{-3}, i.e. $\mathfrak{L}$ is a homogeneous function of the g^{ik} of degree -3. From Euler's theorem we thus get the equation

$$\frac{\partial \mathfrak{L}}{\partial g^{ik}} g^{ik} = -3\mathfrak{L}. \tag{141}$$

In the same way we see that $\mathfrak{L}$ as a function of the variables g_l^{ik} is homogeneous of degree 2. Hence we also have

$$\frac{\partial \mathfrak{L}}{\partial g_l^{ik}} g_l^{ik} = 2\mathfrak{L}. \tag{142}$$

126. The laws of conservation of energy and momentum

In § 115 it was shown that the conservation of electric charge is a consequence of the covariant divergence relation $\operatorname{div}\{s^i\} = 0$. This is connected with the circumstance that the vanishing of a covariant divergence of a four-vector is equivalent to the vanishing of an ordinary divergence of a vector density. By subsequent integration over the space coordinates we are then at once led to the conclusion that the total charge is constant in time.

The law of conservation of energy and momentum, which has the form (X. 41) or

$$\frac{\partial}{\partial x^k}\{\sqrt{(-g)}T_i^k\} = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\sqrt{(-g)}T^{kl} \equiv k_i \tag{143}$$

is, however, not in general equivalent to the vanishing of an ordinary divergence and will therefore not immediately give rise to any conservation laws by integration over the space coordinates. Only in the case of a stationary system considered in § 114 is the right-hand side of (143) zero for $i = 4$, and by subsequent integration over the space coordinates we get a constant of the motion which may be interpreted as the total energy.

The occurrence of the term on the right-hand side of (143) indicates that the system is not strictly closed, this term being analogous to the external four-force density on a non-closed system in the special theory of relativity (cf. Chapter VII). In the case of electromagnetic forces, it was possible by means of Maxwell's equations to write the four-force density as the divergence of the electromagnetic energy-momentum tensor. Similarly, by virtue of the field equations (12), the term

$$k_i = \frac{1}{2}\frac{\partial g_{kl}}{\partial x^i}\sqrt{(-g)}T^{kl} \tag{144}$$

on the right-hand side of (143) can be written in a covariant way in the form

$$k_\iota = -\frac{\partial\{\sqrt{(-g)}t_\iota^{\,k}\}}{\partial x^k}, \tag{145}$$

where $t_\iota^{\,k}$ is a scheme of 4^2 quantities depending on the components of the metric tensor and their first derivatives. The quantities k_ι and $t_\iota^{\,k}$ will of course not transform like tensors, but the equations (143), (144), (145) will, as we shall see, by integrations over the space coordinates, give rise to conservation theorems for quantities which have simple transformation properties and a simple physical meaning.†

In order to prove (145) we first remark that (144) can also be written

$$k_\iota = -\frac{1}{2}\frac{\partial g^{kl}}{\partial x^\iota}\sqrt{(-g)}T_{kl} \tag{144'}$$

by virtue of the rules for lowering indices together with the relations (IX. 7).

Next we substitute from (12) in (144′) and use (140), which gives

$$\begin{aligned} k_\iota = -\tfrac{1}{2}\sqrt{(-g)}T_{kl}\frac{\partial g^{kl}}{\partial x^\iota} &= \frac{1}{2\kappa}\sqrt{(-g)}M_{lm}\frac{\partial g^{lm}}{\partial x^\iota} \\ &= \frac{1}{2\kappa}\left\{\frac{\partial\mathfrak{H}}{\partial g^{lm}}\frac{\partial g^{lm}}{\partial x^\iota}-\frac{\partial}{\partial x^k}\left(\frac{\partial\mathfrak{H}}{\partial g_k^{lm}}\right)\frac{\partial g^{lm}}{\partial x^\iota}\right\} \\ &= \frac{1}{2\kappa}\left\{\frac{\partial\mathfrak{H}}{\partial g^{lm}}\frac{\partial g^{lm}}{\partial x^\iota}+\frac{\partial\mathfrak{H}}{\partial g_k^{lm}}\frac{\partial^2 g^{lm}}{\partial x^\iota\partial x^k}-\frac{\partial}{\partial x^k}\left(\frac{\partial\mathfrak{H}}{\partial g_k^{lm}}\frac{\partial g^{lm}}{\partial x^\iota}\right)\right\}. \end{aligned}$$

As $\mathfrak{H}$ is a function of g^{lm} and g_k^{lm} only and

$$\frac{\partial^2 g^{lm}}{\partial x^\iota\partial x^k} = \frac{\partial g_k^{lm}}{\partial x^\iota},$$

the sum of the first two terms inside the brackets is equal to $\partial\mathfrak{H}/\partial x^\iota$.

Thus we see that k_ι is of the form (145) with

$$\begin{aligned} \sqrt{(-g)}t_\iota^{\,k} &= \frac{1}{2\kappa}\left\{\frac{\partial\mathfrak{H}}{\partial g_k^{lm}}g_\iota^{lm}-\delta_\iota^k\,\mathfrak{H}\right\} \\ &= \frac{1}{2\kappa}\left\{\frac{\partial\mathfrak{L}}{\partial g_k^{lm}}g_\iota^{lm}-\delta_\iota^k\mathfrak{L}-\delta_\iota^k\,2\lambda\sqrt{(-g)}\right\} \end{aligned} \tag{146}$$

Using (145) in (143) the laws of conservation of energy and momentum can be written in the form

$$\frac{\partial\mathfrak{T}_\iota^{\,k}}{\partial x^k} = 0, \tag{147}$$

where

$$\mathfrak{T}_{i\cdot}^{\,k} = \sqrt{(-g)}(T_\iota^k+t_\iota^{\,k}). \tag{148}$$

† A. Einstein, *Berl. Ber.*, p. 778 (1915), *Ann. d Phys.* **49**, 769 (1916), *Berl. Ber.*, p. 448 (1918), F. Klein, *Gött. Nachr., Math.-phys. Klasse*, p 394 (1918).

The quantities $t_i^{\ k}/\sqrt{(-g)}$ do not in general transform like a tensor. Since $\mathfrak{L}$ is quadratic in g_l^{lm}, the quantities $t_i^{\ k}$ can, if we neglect the λ-terms, even be made equal to zero at any given point by introducing a system of coordinates which is geodesic at this point. $\mathfrak{T}_i^{\ k}/\sqrt{(-g)}$ will behave like a tensor only if the transformation coefficients α_i^k are constant. However, by integration over the spatial coordinates we can derive a set of quantities which behave like a four-vector under a much wider group of transformations. Let us consider an isolated system for which the T_i^k are different from zero only inside a certain tube with a finite space-like extension in 4-space. For sufficiently large space-like distances from the tube, we may assume that for a proper choice of coordinates the g_{ik} reduce to the constant values G_{ik} of the special theory of relativity. Such coordinates, which can be quite arbitrary inside the tube, will be called quasi-Galilean. If we neglect the λ-term, the $\mathfrak{T}_i^{\ k}$ will therefore be zero at sufficiently large spatial distances from our system, and by integration of (147) over the space coordinates x^1, x^2, x^3 we get at once

$$\frac{d}{dx^4}\int \mathfrak{T}_i^{\ 4}\, dx^1dx^2dx^3 = 0. \tag{149}$$

Hence we see that the quantities

$$P_i = \frac{1}{c}\int \mathfrak{T}_i^{\ 4}\, dx^1dx^2dx^3 = \frac{1}{c}\int \sqrt{\left(1+\frac{2\chi}{c^2}\right)}\,(T_i^4+t_i^{\ 4})\, dV \tag{150}$$

are constant in time.

Further, it is easily seen that the P_i are invariant for any transformation of coordinates $x'^i = x'^i(x^k)$ inside the tube which leaves the x^i unchanged outside the tube. To prove this we only need to introduce a third system of coordinates x''^i which on the hypersurface Ω_1 defined by $x''^4 = a =$ constant coincides with the system (x^i) and on the different hypersurface Ω_2, where $x''^4 = b$, coincides with the system x'^i. Since the constancy of the quantities (150) holds in all three systems of coordinates we get at once

$$P_i = P_i|_{\Omega_1} = P''_i|_{\Omega_1} = P''_i|_{\Omega_2} = P'_i|_{\Omega_2} = P'_i. \tag{151}$$

On the other hand, for linear orthogonal transformations

$$x'^i = \alpha_k^i\, x^k \tag{152}$$

with a determinant $\alpha = |\alpha_i^k| = \pm 1$, the quantities P_i will transform like the covariant components of a four-vector in a pseudo-Cartesian system of coordinates. For the proof, we first remark that the quantities $\mathfrak{T}_{i\cdot}^{\ k}$ transform like the mixed components of a tensor under the transformations (152) with constant coefficients α_ι^k. Now consider a four-vector a^ι whose components are constant inside the tube. The components

$$a'^\iota = \alpha_k^i a^k \tag{153}$$

are then also constant inside the tube, and the quantities

$$b^k = \mathfrak{T}_\iota{}^k a^\iota \tag{154}$$

will, on account of (147), satisfy the equations

$$\frac{\partial b^k}{\partial x^k} = 0 \tag{155}$$

in all systems of coordinates connected by the linear orthogonal transformations (152). Thus, formally, the situation here is the same as in § 63, apart from the use of the imaginary time representation in Chapter VI. By the same reasoning as that used for the derivation of (VI. 31), it now follows from (155) that the quantity $(b^4/c)\,dx^1dx^2dx^3 = P_\iota\, a^\iota$ is invariant under the orthogonal transformations (152) and, as this must hold for an arbitrary constant vector a^ι, the integrals P_ι must transform like the covariant components of a four-vector. Combining this result with the fact expressed by (151) we see that the P_ι must behave like a four-vector by any transformation which outside the tube has the form of a Lorentz transformation.

The quantities $P_\iota = (P_\iota, -E/c)$ represent the total momentum **P** and energy E of an isolated system. If $t_\iota{}^4 = 0$, the expression for the total energy E reduces to the expression obtained in § 114 for a stationary system. The last term in the integrals (150) may be interpreted as the contribution of the gravitational field to the total momentum and energy. As already mentioned in § 112, a unique separation into a material and a gravitational part is not, however, possible. The separation depends, in general, on the coordinates used in the evaluation of the energy and momentum.

It is now tempting to consider (147) as the differential form of the conservation laws. The quantities

$$\left.\begin{aligned} S^\iota &= -c\sqrt{\left(1+\frac{2\chi}{c^2}\right)}(T_4^\iota+t_4{}^\iota) \\ h &= -\sqrt{\left(1+\frac{2\chi}{c^2}\right)}(T_4^4+t_4{}^4) \\ g_\iota &= \frac{\sqrt{(1+2\chi/c^2)}}{c}(T_\iota^4+t_\iota{}^4) \end{aligned}\right\}, \tag{156}$$

which in the case of $t_\iota{}^k = 0$ reduce to the expressions (X. 45, 47), will then be interpreted as the energy flux, energy density, and momentum density, respectively. It should be remarked, however, that this interpretation is not independent of the coordinates used; for t_4^{ι} and t_{ι}^{4} and therefore S^{ι} and g_{ι} will not in general transform like the components of space vectors by a simple change of the spatial coordinates, neither will h in general be an invariant under such transformations. Only the expressions (150) for the *total* momentum and energy have an invariant meaning which is physically satisfactory.

127. Different expressions for the densities of energy and momentum

By means of (146) and (132) we can now write down explicit expressions for the quantities t_{ι}^{k}. As shown in Appendix 9, the derivative of $\mathfrak{L}$ with respect to g_k^{lm} is given by

$$\frac{\partial \mathfrak{L}}{\partial g_k^{lm}} = \sqrt{(-g)}\{\Gamma_{lm}^{k}-\tfrac{1}{2}(\delta_l^k\,\Gamma_{mr}^{r}+\delta_m^k\,\Gamma_{lr}^{r})-\tfrac{1}{2}(g^{rs}\Gamma_{rs}^{k}-g^{rk}\Gamma_{rs}^{s})g_{lm}\}. \quad (157)$$

After a simple calculation we then get, remembering that

$$\Gamma_{rs}^{s} = \frac{1}{\sqrt{(-g)}}\frac{\partial\sqrt{(-g)}}{\partial x^r} = -\tfrac{1}{2}g_{lm}\frac{\partial g^{lm}}{\partial x^r},$$

$$\sqrt{(-g)}t_{i\cdot}^{\;k} = \frac{1}{2\kappa}\left\{\Gamma_{lm}^{k}\frac{\partial\{\sqrt{(-g)}g^{lm}\}}{\partial x^i}-\Gamma_{ms}^{s}\frac{\partial\{\sqrt{(-g)}g^{km}\}}{\partial x^i}-\delta_i^k\mathfrak{H}\right\}. \quad (158)$$

As shown by Tolman,† the quantities $\mathfrak{T}_{\iota}^{\;k}$ may also be expressed in the very useful form of an ordinary divergence

$$\mathfrak{T}_{i\cdot}^{\;k} = \frac{\partial s_i^{km}}{\partial x^m}. \quad (159)$$

From (148), (13), and (146) we get

$$\mathfrak{T}_{\iota}^{\;k} = \sqrt{(-g)}(T_i^k+t_{i\cdot}^{\;k}) = -\frac{\sqrt{(-g)}}{\kappa}M_i^k+\sqrt{(-g)}t_{i\cdot}^{\;k}$$

$$= \frac{1}{\kappa}\left\{-\sqrt{(-g)}(R_i^k-\tfrac{1}{2}Rg_i^k)+\frac{1}{2}\frac{\partial\mathfrak{L}}{\partial g_k^{lm}}g_i^{lm}-\tfrac{1}{2}\delta_i^k\mathfrak{L}\right\} \quad (160)$$

in which the λ-term has disappeared. This expression can be further transformed by means of (139):

$$\mathfrak{T}_{i\cdot}^{\;k} = \frac{1}{\kappa}\left\{g^{kl}\left[\frac{\partial}{\partial x^m}\left(\frac{\partial\mathfrak{L}}{\partial g_m^{il}}\right)-\frac{\partial\mathfrak{L}}{\partial g^{il}}\right]+\frac{1}{2}\frac{\partial\mathfrak{L}}{\partial g_k^{lm}}g_i^{lm}-\tfrac{1}{2}\delta_i^k\mathfrak{L}\right\}$$

$$= \frac{1}{\kappa}\frac{\partial}{\partial x^m}\left[g^{kl}\left(\frac{\partial\mathfrak{L}}{\partial g_m^{il}}\right)\right]+\frac{1}{\kappa}\left[\frac{1}{2}\frac{\partial\mathfrak{L}}{\partial g_k^{lm}}g_i^{lm}-\frac{\partial\mathfrak{L}}{\partial g_m^{il}}g_m^{kl}-g^{kl}\frac{\partial\mathfrak{L}}{\partial g^{il}}-\tfrac{1}{2}\delta_i^k\mathfrak{L}\right]. \quad (161)$$

† R. C. Tolman, *Phys. Rev.* **35**, 875 (1930).

As shown in Appendix 9, the last term is, however, identically zero. Hence, $\mathfrak{T}_i^k$ may be written in the form (159) of an ordinary divergence of the quantity

$$s_i^{km} = \frac{1}{\kappa}\frac{\partial \mathfrak{L}}{\partial g_m^{il}} g^{kl}. \tag{162}$$

Further, we get from (148), (146), (142), and (159), neglecting the small λ-terms,

$$\mathfrak{T}_i^i = \sqrt{(-g)}T_i^i + \sqrt{(-g)}t_i^i = \sqrt{(-g)}T_i^i + \frac{1}{2\kappa}\left(\frac{\partial \mathfrak{L}}{\partial g_i^{lm}} g_i^{lm} - 4\mathfrak{L}\right)$$
$$= \sqrt{(-g)}T_i^i - \frac{\mathfrak{L}}{\kappa} = \frac{\partial s_i^{im}}{\partial x^m}. \tag{163}$$

Substituting the expression for $\mathfrak{L}$ obtained from this equation in the expression (146) we get for t_4^4 and, finally, for the 'energy density'

$$\mathfrak{T}_4^4 = \sqrt{(-g)}T_4^4 + \sqrt{(-g)}t_4^4 = \sqrt{(-g)}\tfrac{1}{2}(T_4^4 - T_1^1 - T_2^2 - T_3^3) +$$
$$+ \frac{1}{2}\frac{\partial s_l^{lm}}{\partial x^m} + \frac{1}{2\kappa}\frac{\partial \mathfrak{L}}{\partial g_4^{lm}}\frac{\partial g^{lm}}{\partial x^4}, \tag{164}$$

from which the total energy is obtained by integration over the spatial coordinates.

128. The gravitational mass and total energy and momentum of an isolated system

By an isolated system we understand a system which allows the introduction of quasi-Galilean coordinates $x^i = (x, y, z, ct)$ in which the line element at great distance from the system takes the form (86′), where m is a constant. For the total energy-momentum vector P_i of the system we then get by means of (150) and (159)

$$P_i = \frac{1}{c}\int \mathfrak{T}_i^4\, dx^1dx^2dx^3 = \frac{1}{c}\int \frac{\partial s_i^{4\mu}}{\partial x^\mu}\, dx^1dx^2dx^3 +$$
$$+ \frac{1}{c^2}\frac{\partial}{\partial t}\int s_i^{44}\, dx^1dx^2dx^3. \tag{165}$$

The first integral on the right-hand side can be transformed into an integral over an infinitely distant spherical surface

$$r = (x^2+y^2+z^2)^{\frac{1}{2}} = \text{constant}.$$

It thus depends only on the values of the metric tensor and its derivatives at large distances which are time-independent. As the vector P_i is constant for an isolated system, the last integral must also be constant;

actually it must have the value zero, since the integral involved cannot change permanently at a constant finite rate in an isolated system. Hence, we get

$$P_i = \frac{1}{c}\int s_i^{4\mu} n_\mu \, d\sigma, \tag{166}$$

where $n_\mu = \partial r/\partial x^\mu = (x/r, y/r, z/r)$ is a unit vector in the direction of the outward normal, and the integral is extended over the surface of the infinitely distant sphere $r =$ constant. At large distances we have, however, according to (86′),

$$g_{11} = g_{22} = g_{33} = 1+\frac{2m}{r}, \qquad g_{44} = -\left(1-\frac{2m}{r}\right),$$

all other components being zero. Hence,

$$\frac{\partial g_{ik}}{\partial x^l} = \delta_{ik}\, 2m \frac{\partial(1/r)}{\partial x^l} = -\delta_{ik}(1-\delta_{l4})\frac{2m}{r^2} n_l. \tag{167}$$

If we neglect terms of the order m/r compared with unity, we can replace g^{ik} by G^{ik} and $\sqrt{(-g)}$ by 1 when multiplied by a Christoffel symbol Γ^i_{kl}. From (162) and (157) we therefore get to this approximation

$$s_i^{4\mu} = \frac{1}{\kappa}\{\Gamma^\mu_{il} - \tfrac{1}{2}(\delta^\mu_i \Gamma^r_{lr} + \delta^\mu_l \Gamma^r_{ir}) - \tfrac{1}{2}(G^{rs}\Gamma^\mu_{rs} - G^{r\mu}\Gamma^s_{rs})G_{il}\}G^{4l}. \tag{168}$$

As $G^{ik} = 0$ for $i \neq k$ the only Christoffel symbols occurring in this expression are the following, which are easily calculated by means of (IX. 50) and (167):

$$\Gamma^\mu_{i4} = \delta^4_i \frac{m}{r^2} n_\mu, \qquad \Gamma^s_{rs} = -(1-\delta_{r4})\frac{2m}{r^2} n_r, \qquad G^{rs}\Gamma^\mu_{rs} = G^{rr}\Gamma^\mu_{rr} = 0. \tag{169}$$

Hence we get

$$s_i^{4\mu} = -\delta^4_i \frac{2m}{\kappa r^2} n_\mu, \tag{170}$$

and from (166)

$$P_i = -\delta^4_i \frac{1}{c}\int \frac{2m}{\kappa r^2}\left[\left(\frac{x}{r}\right)^2 + \left(\frac{y}{r}\right)^2 + \left(\frac{z}{r}\right)^2\right] d\sigma = -\delta^4_i \frac{8\pi m}{\kappa c}. \tag{171}$$

As $P_i = (\mathbf{P}, -E/c)$, we see that the total momentum is zero in the system of coordinates corresponding to the line element (86′) at infinity, the isolated system is as a whole at rest in this system. For the total energy we have, on account of (88),

$$E = -cP_4 = \frac{8\pi m}{\kappa} = Mc^2, \tag{172}$$

i.e. the gravitational mass M, as determined by the scalar gravitational potential at great distances, is connected with the total energy by

Einstein's relation. For a Lorentz transformation of the quasi-Galilean coordinates the quantities P_ι transform like a four-vector and in the transformed system **P** and E will have the same values as the momentum and energy of a moving particle in the special theory of relativity.

In the calculation of the total energy just worked out we have used the formula (159) which expresses the energy as a function of pure gravitational field quantities, actually only the field variables at infinity occurring in the expression (166). Sometimes it is convenient to use instead the formula (164). For a stationary or a quasi-stationary system where all terms depending on time derivatives can be neglected, we get

$$E = -cP_4 = -\int \mathfrak{T}_4{}^4\, dx^1dx^2dx^3$$

$$= -\int \tfrac{1}{2}\sqrt{(-g)}\,(T_4^4-T_1^1-T_2^2-T_3^3)\,dx^1dx^2dx^3-\tfrac{1}{2}\int s_\iota^{\iota\mu}n_\mu\,d\sigma. \tag{173}$$

In the last integral the integration is again extended over the distant sphere of radius $r =$ constant, hence we get by means of (162), (157), (169), and (170)

$$s_\iota^{\iota\mu} = \frac{1}{\kappa}\{\Gamma^\mu_{\iota l}-\tfrac{1}{2}(\delta^\mu_\iota\,\Gamma^r_{lr}+\delta^\mu_l\,\Gamma^r_{\iota r})-\tfrac{1}{2}(G^{rs}\Gamma^\mu_{rs}-G^{r\mu}\Gamma^s_{rs})G_{\iota l}\}G^{\iota l}$$

$$= -\frac{2m}{\kappa r^2}n_\mu = s_4^{4\mu},$$

and a comparison with (166) shows that the last integral in (173) is equal to $-E$. Thus we get for E the simple expression

$$E = -\int \sqrt{(-g)}(T_4^4-T_1^1-T_2^2-T_3^3)\,dx^1dx^2dx^3$$

$$= \int \sqrt{\left(1+\frac{2\chi}{c^2}\right)}(T_1^1+T_2^2+T_3^3-T_4^4)\,dV, \tag{174}$$

which contains only an integration over the part of space where T_ι^k is different from zero, i.e. where there is actually matter present. A comparison of (174) with (172) shows that the quantity

$$\mu = \frac{\sqrt{(1+2\chi/c^2)}}{c^2}(T_1^1+T_2^2+T_3^3-T_4^4) \tag{175}$$

may be interpreted as the density of gravitational mass in the system. Since μ obviously behaves like a scalar for all purely spatial transformations this interpretation has a well-defined physical meaning.

In the case of an incompressible fluid at rest, treated in § 124, we have, according to (92′),

$$T_1^1 = T_2^2 = T_3^3 = \overset{0}{p}, \qquad T_4^4 = -\overset{0}{\mu}{}^0c^2.$$

Further, for $\lambda = 0$, we have on account of (103), (104), (102), (107), and (175),

$$\sqrt{(1+2\chi/c^2)} = A - B\sqrt{(1-r^2/R^2)}$$

$$\sqrt{\gamma} = \frac{r^2 \sin\theta}{\sqrt{(1-r^2/R^2)}}, \qquad dV = \sqrt{\gamma}\, dr d\theta d\phi = \frac{r^2 \sin\theta\, dr d\theta d\phi}{\sqrt{(1-r^2/R^2)}},$$

$$\mu = (A - B\sqrt{(1-r^2/R^2)})(3\mathring{p}/c^2 + \mathring{\mu}^0) = \mathring{\mu}^0\sqrt{(1-r^2/R^2)}. \tag{176}$$

Hence, we get for the total field-producing gravitational mass of the spherical fluid

$$M = \int \mu\, dV = 4\pi \int\limits_0^{r_1} r^2 \mathring{\mu}^0\, dr = \frac{4\pi r_1^3}{3} \mathring{\mu}^0, \tag{177}$$

in accordance with (108). While the proper rest mass density $\mathring{\mu}^0$ is constant in an incompressible fluid, the gravitational mass density μ decreases with r according to (176).

XII

EXPERIMENTAL VERIFICATION OF THE GENERAL THEORY OF RELATIVITY. COSMOLOGICAL PROBLEMS

129. The gravitational shift of spectral lines

WHILE the consequences of the special theory of relativity have been verified to a very high degree of accuracy by numerous experiments, the experimental verification of the general theory has so far been limited to three cases only. The reason for this is obvious and is connected with the fact that the Newtonian gravitational theory represents a very good approximation for all gravitational phenomena inside the solar system.

The most elementary of the effects which represent a crucial test of the general theory of relativity is the gravitational shift of spectral lines, which is a direct consequence of the principle of equivalence. According to (VIII. 100), the rate of a standard clock at rest in a gravitational field at a place with the scalar gravitational potential χ is connected with the rate of the coordinate clock defining the time t of our system of coordinates by the formula†

$$d\tau_0 = dt\,(1+2\chi/c^2)^{\frac{1}{2}} \tag{1}$$

Now consider an atom at rest at the place p_1 with the scalar potential χ_1, emitting light of the proper frequency $\overset{0}{\nu}$. Then $\overset{0}{\nu}$ is equal to the number of light waves emitted per unit time in the time-scale of a local rest system of inertia which is the same as the scale of the standard clock at rest at the point p_1. The number of waves emitted per unit time in the scale of the corresponding coordinate clock will then be

$$\nu_1 = \overset{0}{\nu}(1+2\chi_1/c^2)^{\frac{1}{2}}, \tag{2}$$

for, according to (1), the time-interval $\Delta\tau$ of the standard clock corresponding to $\Delta t = 1$ is $\Delta\tau = (1+2\chi_1/c^2)^{\frac{1}{2}}$. If the gravitational field is stationary, the g_{ik} are independent of t and the number of waves reaching an observer at an arbitrary point p_2 per unit time in the t-scale must be constant in time and equal to the number of waves ν_1 emitted from p_1

† This formula, which is a consequence of the principle of equivalence, has also played an important part in discussions between Einstein and Bohr on the consistency of the quantum mechanical description See the article by N Bohr in *Albert Einstein Philosopher-Scientist*, Library of Living Philosophers, vol vii, Evanston, 1949.

per unit time in the same scale. The frequency ν of the radiation measured by a standard clock at rest at p_2 will then be

$$\nu = \nu_1(1+2\chi_2/c^2)^{-\frac{1}{2}}, \tag{3}$$

where χ_2 is the scalar potential at the place p_2, for a unit time-interval in the scale of the standard clock at p_2 corresponds to a time-interval

$$\Delta t = (1+2\chi_2/c^2)^{-\frac{1}{2}}.$$

From (2) and (3) we thus get

$$\nu = \overset{0}{\nu}\left(\frac{1+2\chi_1/c^2}{1+2\chi_2/c^2}\right)^{\frac{1}{2}}. \tag{4}$$

Hence the observed frequency ν will differ from the proper frequency $\overset{0}{\nu}$ by an amount $\Delta\nu = \nu-\overset{0}{\nu}$, which in the case of weak fields is given by

$$\frac{\Delta\nu}{\overset{0}{\nu}} = \frac{\nu-\overset{0}{\nu}}{\nu^0} = \frac{\Delta\chi}{c^2}, \tag{5}$$

where $\Delta\chi = \chi_1-\chi_2$ is the difference between the potentials at the places where the light is emitted and observed, respectively.

In the gravitational field of the sun we have, according to (XI. 83, 88),

$$\chi = -\frac{mc^2}{r} = -\frac{kM}{r},$$

where

$$M = 1{\cdot}983\times 10^{33}\ \text{gm.} \tag{5'}$$

is the mass of the sun. Hence we get for the shift of a spectral line emitted by an atom in the outer layers of the sun as compared with the same line emitted on the earth

$$\frac{\Delta\nu}{\overset{0}{\nu}} = -\frac{kM}{c^2}\left(\frac{1}{r_1}-\frac{1}{r_2}\right), \tag{6}$$

where r_1 and r_2 are the r-values at the surface of the sun and at the distance of the earth, respectively. These values of r are not exactly equal to the values of the distances calculated by means of the usual astronomical methods which are based on the assumption that the space is Euclidean and that the light rays are moving in straight lines. The differences are small, however, and since the effect (6) is itself small, the error introduced by using the astronomical values of the distances instead of the correct r-values is small of a higher order. Since $r_2 \gg r_1$, we get, putting r_1 equal to the usual astronomical value of the radius of the sun,

$$r_1 = 6{\cdot}95\times 10^{10}\ \text{cm.}, \tag{6'}$$

the value

$$\frac{\Delta\nu}{\overset{0}{\nu}} = -2{\cdot}12\times 10^{-6} \tag{7}$$

for the line shift. The observed spectral lines should therefore be shifted slightly towards the red. This interesting effect predicted by Einstein† is in satisfactory agreement with the observations both in the case of the sun and in the case of the heavy companion of Sirius, where the effect is about thirty times larger.‡

130. The advance of the perihelion of Mercury

Let us consider the motion of a particle (a planet) in the gravitational field of a much heavier body (the sun) described by Schwarzschild's exterior solution (XI. 83, 88, 24). In this case, the components of the metric tensor are

$$\left.\begin{aligned} g_{11} &= \frac{1}{1-\alpha/r}, \qquad g_{22} = r^2, \qquad g_{33} = r^2\sin^2\theta, \qquad g_{44} = -\left(1-\frac{\alpha}{r}\right) \\ g_{ik} &= 0 \text{ for } i \neq k, \qquad \alpha = \frac{2kM}{c^2} = \frac{\kappa Mc^2}{4\pi} \end{aligned}\right\}. \tag{8}$$

Hence

$$\gamma_\iota = 0, \qquad \gamma_{\iota\kappa} = g_{\iota\kappa}, \qquad \chi = -\frac{c^2}{2}(1+g_{44}) = -\frac{\alpha c^2}{2r} = -\frac{kM}{r}. \tag{9}$$

According to (X. 17, 17′), the gravitational force on the particle is

$$\left.\begin{aligned} K_\iota &= \left(-\frac{m\alpha c^2}{2r^2}, 0, 0\right) \\ \text{where} \qquad m &= \mathring{m}_0\,\Gamma = \frac{\mathring{m}_0}{\sqrt{(1-\alpha/r-u^2/c^2)}} \end{aligned}\right\} \tag{10}$$

is the relativistic mass of the planet. The square of the velocity is

$$u^2 = \gamma_{\iota\kappa}u^\iota u^\kappa = \frac{\dot{r}^2}{1-\alpha/r}+r^2\dot{\theta}^2+r^2\sin^2\theta\,\dot{\phi}^2, \tag{11}$$

where the dots mean differentiation with respect to the time-variable t. Further, by (X. 12) and (8), the momentum vector of the particle is

$$p_\iota = \mathring{m}_0\,\Gamma\left(\frac{\dot{r}}{1-\alpha/r}, r^2\dot{\theta}, r^2\sin^2\theta\,\dot{\phi}\right). \tag{12}$$

The equations of motion (X. 15)

$$\frac{d_c p_\iota}{dt} \equiv \frac{dp_\iota}{dt}-\frac{1}{2}\frac{\partial\gamma_{\kappa\lambda}}{\partial x^\iota}u^\kappa p^\lambda = K_\iota \tag{13}$$

† A. Einstein, *Ann. d. Phys.* **35**, 898 (1911).

‡ St. John, *Astrophys. Journ.* **67**, 195 (1928); W. Adams, *Proc. Nat. Acad.* **11**, 382 (1925); see also E. F. Freundlich and W. Ledermann, *Mon. Not. Roy. Astr. Soc.* **104**, 1, 40 (1944).

are thus in the present case

$$\frac{d}{dt}\left(\frac{\Gamma\dot{r}}{1-\alpha/r}\right)-\frac{1}{2}\left[-\frac{(\alpha/r^2)\Gamma}{(1-\alpha/r)^2}\dot{r}^2+2r\Gamma\dot{\theta}^2+2r\sin^2\theta\,\Gamma\dot{\phi}^2\right] = -\Gamma\frac{\alpha c^2}{2r^2}, \quad (14a)$$

$$\frac{d}{dt}(\Gamma r^2\dot{\theta})-r^2\sin\theta\cos\theta\,\Gamma\dot{\phi}^2 = 0, \quad (14b)$$

and

$$\frac{d}{dt}(\Gamma r^2\sin^2\theta\,\dot{\phi}) = 0. \quad (14c)$$

Since the gravitational field is static, the energy H of the particle defined by (X. 22) is constant. Hence

$$\frac{H}{\mathring{m}_0 c^2} \equiv \Gamma\left(1+\frac{2\chi}{c^2}\right) \equiv \Gamma\left(1-\frac{\alpha}{r}\right) = E \quad (15)$$

is a first integral of the equations of motion, where E is a constant of integration representing the energy of the system divided by $\mathring{m}_0 c^2$.

From (14 b) we see that

$$\theta = \tfrac{1}{2}\pi \quad (16)$$

represents another integral of the equations of motion. On account of the central symmetry of our problem, any plane through the centre may, however, be chosen as the plane $\theta = \frac{1}{2}\pi$, i.e. the orbit can lie in any plane through the centre. Substituting from (16) in (14 c) we get at once the further integral

$$\Gamma r^2\dot{\phi} = CE, \quad (17)$$

where C is another constant of integration. On account of (15) this equation can also be written

$$\frac{r^2\dot{\phi}}{1-\alpha/r} = C. \quad (18)$$

Now even for Mercury, the planet nearest to the sun, the quantity α/r is a very small quantity of the order of magnitude

$$\frac{\alpha}{r} = \frac{2kM}{c^2 r} \approx 5\times 10^{-8}, \quad (19)$$

and for the other planets α/r is much smaller. Therefore, in all actual cases, the gravitational field may be treated as a weak field and, since also $u^2/c^2 \ll 1$, we may to a first approximation use the approximation (X. 25) for H in (15). Further, using (11) and (16) we therefore get for (15) to this approximation

$$\tfrac{1}{2}(\dot{r}^2+r^2\dot{\phi}^2)-\frac{kM}{r} = \text{constant}, \quad (20)$$

which is the usual energy equation in Newton's theory. To the same approximation, (18) reduces to

$$r^2\dot{\phi} = C, \tag{21}$$

expressing the conservation of angular momentum. The orbit determined by (20) and (21) is therefore an ellipse

$$r = \frac{a}{(1+\epsilon\cos(\phi-\phi_0))}, \tag{22}$$

where ϵ is the eccentricity, and

$$r_1 = \frac{a}{(1-\epsilon)}, \qquad r_2 = \frac{a}{(1+\epsilon)} \tag{23}$$

are the r-values corresponding to aphelion and perihelion of the planet, respectively. For Mercury we have

$$\epsilon = 0{\cdot}2056, \qquad a = 5{\cdot}786\times 10^{12}\ \text{cm}. \tag{24}$$

For stronger fields the equation (20) has to be replaced by the energy equation (15), and instead of (21) we have the integral (18). However, the left-hand side of (18) cannot in general be interpreted as angular momentum, since the notion of a 'radius vector' occurring in the definition of the angular momentum has an unambiguous meaning only in a Euclidean space.

In order to determine the orbit in the general case, we introduce the quantity $\rho = 1/r$ instead of r. Thus we get from (18)

$$\dot{r} = \frac{dr}{d\phi}\dot{\phi} = \frac{dr}{d\rho}\frac{d\rho}{d\phi}\frac{C}{r^2}\left(1-\frac{\alpha}{r}\right) = -\left(\frac{d\rho}{d\phi}\right)C(1-\alpha\rho), \tag{25}$$

and by means of (11), (16), (18), and (25) we get for the velocity of the particle

$$u^2 = C^2(1-\alpha\rho)\left[\left(\frac{d\rho}{d\phi}\right)^2+\rho^2-\alpha\rho^3\right]. \tag{26}$$

Using the expression (10) for Γ and squaring the energy equation (15) we get

$$(1-\alpha\rho)^2 = E^2\left[1-\alpha\rho-\frac{u^2}{c^2}\right]. \tag{27}$$

Thus, substituting the expression (26) for u^2 and solving with respect to $(d\rho/d\phi)^2$, we get the differential equation for the orbit of the particle in the form

$$\left(\frac{d\rho}{d\phi}\right)^2 = A+B\rho-\rho^2+\alpha\rho^3, \tag{28}$$

where A and B are constants. On account of (19) the last term in this equation is very small compared with the term ρ^2 and, if we neglect it entirely, (28) reduces to the equation obtained from the Newtonian equations (20), (21), i e.

$$\left(\frac{d\rho}{d\phi}\right)^2 = A+B\rho-\rho^2. \tag{29}$$

It is easily seen that (22) is a solution of this approximate equation. The maximum and minimum values of ρ corresponding to perihelion and aphelion are obtained as roots of the quadratic equation

$$A+B\rho-\rho^2 = 0. \tag{30}$$

Let ρ_1 and ρ_2 be the two roots which must be real and positive in the case of an orbit which stays inside a finite region in space. Then we have, according to (23),

$$\rho_1 = \frac{1-\epsilon}{a}, \qquad \rho_2 = \frac{1+\epsilon}{a}. \tag{31}$$

The equation (29) may now be written

$$\frac{d\rho}{d\phi} = \pm\sqrt{\{(\rho-\rho_1)(\rho_2-\rho)\}}, \tag{32}$$

and the increase $\phi_2-\phi_1$ in ϕ during an increase of ρ from ρ_1 to ρ_2 is obtained by integration

$$\phi_2-\phi_1 = \int\limits_{\rho_1}^{\rho_2} \frac{d\rho}{\sqrt{\{(\rho-\rho_1)(\rho_2-\rho)\}}} = \sin^{-1}\frac{\rho-\frac{1}{2}(\rho_1+\rho_2)}{\frac{1}{2}(\rho_2-\rho_1)}\Bigg|_{\rho_1}^{\rho_2} = \pi, \tag{33}$$

in accordance with the solution (22).

Returning now to the exact equation (28), we see that the right-hand side of (28) is a polynomial of third degree and the equation obtained by putting $d\rho/d\phi = 0$ has three roots ρ_1, ρ_2, ρ_3. For small values of α two of these roots, ρ_1 and ρ_2 say, must be approximately equal to the two roots (31) of the equation (30), and since

$$\rho_1+\rho_2+\rho_3 = \frac{1}{\alpha}, \tag{34}$$

ρ_3 must be very large for small values of α. The roots ρ_1 and ρ_2 will therefore again represent the minimum and maximum values of ρ in the orbit of the planet. Instead of (32) we now have

$$\begin{aligned}\frac{d\rho}{d\phi} &= \pm\sqrt{\{\alpha(\rho-\rho_1)(\rho-\rho_2)(\rho-\rho_3)\}} \\ &= \pm\sqrt{\left\{(\rho-\rho_1)(\rho_2-\rho)[1-\alpha(\rho_1+\rho_2)]\left[1-\frac{\alpha\rho}{1-\alpha(\rho_1+\rho_2)}\right]\right\}},\end{aligned} \tag{35}$$

where we have used the equation

$$\alpha\rho_3 = 1-\alpha(\rho_1+\rho_2)$$

following from (34). Since $\alpha\rho$ and $\alpha(\rho_1+\rho_2)$ are small quantities we thus get to a first approximation in α

$$d\phi = \pm\frac{d\rho}{\sqrt{\{(\rho-\rho_1)(\rho_2-\rho)\}}}\left[1+\frac{\alpha(\rho_1+\rho_2)}{2}\right]\left[1+\frac{\alpha\rho}{2}\right]$$

and the increase in ϕ for an increase of ρ from ρ_1 to ρ_2 is

$$\phi_2-\phi_1 = \left[1+\frac{\alpha(\rho_1+\rho_2)}{2}\right]\int\limits_{\rho_1}^{\rho_2}\frac{1+\alpha\rho/2}{\sqrt{\{(\rho-\rho_1)(\rho_2-\rho)\}}}\,d\rho$$

$$= \left[1+\frac{\alpha(\rho_1+\rho_2)}{2}\right]\times$$

$$\times\left|-\frac{\alpha}{2}\sqrt{\{(\rho-\rho_1)(\rho_2-\rho)\}}+\left(1+\frac{\alpha}{4}(\rho_1+\rho_2)\right)\sin^{-1}\frac{\rho-\frac{1}{2}(\rho_1+\rho_2)}{\frac{1}{2}(\rho_2-\rho_1)}\right|_{\rho_1}^{\rho_2}.$$

Hence the difference in ϕ between two successive perihelions is

$$2(\phi_2-\phi_1) = 2\left(1+\frac{\alpha(\rho_1+\rho_2)}{2}\right)\left(1+\frac{\alpha}{4}(\rho_1+\rho_2)\right)\pi$$

$$= 2\pi\left[1+\frac{3\alpha}{4}(\rho_1+\rho_2)\right], \tag{36}$$

which differs from the value 2π obtained from (33) by the amount

$$\Delta\phi = \frac{3\pi\alpha}{2}(\rho_1+\rho_2). \tag{37}$$

This result may be interpreted by assigning to each revolution of the planet an advance of its perihelion of the amount (37). Since it is a small effect we can substitute the approximate values (31) for ρ_1 and ρ_2. For Mercury we get, using (24), an advance of the perihelion of an angle 42·9″ per century, which is in satisfactory agreement with the observed advance after subtraction of the effect due to the perturbation by the other planets.† For the other planets the advance of perihelion following from Einstein's theory of gravitation is too small to be observed with certainty.

Instead of the Schwarzschild form (XI. 83) of the line element we could in these calculations also have used the form (XI. 86) corresponding to the 'isotropic' coordinates (XI. 85). In these coordinates the equations of motion are somewhat more complicated, but the final result regarding the advance of perihelion is, of course, the same as that given by (37).

† J. Chazy, *C.R.* **182**, 1134 (1926).

131. The gravitational deflexion of light

The third of the effects which furnishes a crucial test of the general theory of relativity is the deflexion of a light ray in the gravitational field of the sun. Since this field is static, the trajectory of a light ray is, according to the considerations of § 117, determined by Fermat's principle, i.e. by the equations (X. 100, 99, 94):

$$\frac{d_c u_\iota}{d\lambda} = \frac{d}{d\lambda}(g_{\iota\kappa}\dot{x}^\kappa) - \frac{1}{2}\frac{\partial g_{\kappa\lambda}}{\partial x^\iota}\dot{x}^\kappa\dot{x}^\lambda = -\frac{1}{w^3}\frac{\partial w}{\partial x^\iota}, \tag{38}$$

$$\dot{x}^\iota = \frac{dx^\iota}{d\lambda}, \qquad w = c\Big/\sqrt{\left(1+\frac{2\chi}{c^2}\right)}, \qquad \frac{d\sigma}{d\lambda}w = 1. \tag{39}$$

In the field described by (8) and (9) we thus have

$$w = c\Big/\sqrt{\left(1-\frac{\alpha}{r}\right)}, \tag{40}$$

$$\left(\frac{d\sigma}{d\lambda}\right)^2\frac{w^2}{c^2} = \dot{r}^2 + (r^2\dot\theta^2 + r^2\sin^2\theta\,\dot\phi^2)\left(1-\frac{\alpha}{r}\right) = \frac{1}{c^2}, \tag{41}$$

and the two equations (38) with $\iota = 2, 3$ reduce to

$$\frac{d}{d\lambda}(r^2\dot\theta) - \tfrac{1}{2}.2r^2\sin\theta\cos\theta\,\dot\phi^2 = 0, \qquad \frac{d}{d\lambda}(r^2\sin^2\theta\,\dot\phi) = 0. \tag{42}$$

From (42) we see again that

$$\theta = \tfrac{1}{2}\pi, \qquad r^2\dot\phi = C \tag{43}$$

are integrals and (41) then reduces to

$$\dot{r}^2 + r^2\dot\phi^2\left(1-\frac{\alpha}{r}\right) = \dot{r}^2 + \frac{C^2}{r^2} - \frac{C^2\alpha}{r^3} = \frac{1}{c^2}. \tag{44}$$

Introducing $\rho = 1/r$ as a new variable in (44) we get, since

$$\dot{r} = \frac{dr}{d\lambda} = \frac{dr}{d\phi}\frac{d\phi}{d\lambda} = -\frac{1}{\rho^2}\left(\frac{d\rho}{d\phi}\right)C\rho^2 = -C\frac{d\rho}{d\phi},$$

$$\left(\frac{d\rho}{d\phi}\right)^2 = \frac{1}{c^2C^2} - \rho^2 + \alpha\rho^3 = \frac{1}{\Delta^2} - \rho^2 + \alpha\rho^3, \tag{45}$$

where we have put $\Delta = cC$

Now consider a light ray coming from infinity ($\rho = 0$) along the direction $\phi = 0$. Neglecting first the small term $\alpha\rho^3$ in (45), this equation is easily integrated. We get

$$\phi = \int_0^\rho \frac{\Delta\,d\rho}{\sqrt{\{1-(\Delta\rho)^2\}}} = \sin^{-1}(\Delta\rho)\Big|_0^\rho$$

i.e.

$$\rho = \frac{1}{\Delta}\sin\phi, \qquad r = \frac{\Delta}{\sin\phi}. \tag{46}$$

If we make a picture of the trajectory in a Euclidean plane, where the coordinates r, ϕ are pictured as ordinary polar coordinates, the curve (46) represents a straight line which passes the centre $r = 0$ at a distance Δ for $\phi = \frac{1}{2}\pi$ and which goes to infinity again for $\phi \to \pi$.

The exact equation (45) may now be written

$$\frac{d\rho}{d\phi} = \left\{\frac{1}{\Delta^2}-\rho^2+\alpha\rho^3\right\}^{\frac{1}{2}} = \frac{1}{\Delta}(1-\sigma^2)^{\frac{1}{2}}, \tag{47}$$

where we have introduced a new variable

$$\sigma = \Delta\rho(1-\alpha\rho)^{\frac{1}{2}} \tag{48}$$

instead of ρ. Since $\alpha\rho$ is a small quantity, we may write to the first order in α

$$\sigma = \Delta\rho\left(1-\frac{\alpha\rho}{2}\right), \qquad \Delta\rho = \sigma\left(1+\frac{\alpha\rho}{2}\right) = \sigma\left(1+\frac{\alpha\sigma}{2\Delta}\right). \tag{49}$$

Hence

$$\Delta\, d\rho = d\sigma\left(1+\frac{\alpha\sigma}{\Delta}\right),$$

and from (47)

$$\phi = \int\limits_0^\rho \frac{\Delta\, d\rho}{(1-\sigma^2)^{\frac{1}{2}}} = \int\limits_0^\sigma \frac{d\sigma\,(1+\alpha\sigma/\Delta)}{(1-\sigma^2)^{\frac{1}{2}}} = \sin^{-1}\sigma-\frac{\alpha}{\Delta}(1-\sigma^2)^{\frac{1}{2}}+\frac{\alpha}{\Delta}. \tag{50}$$

According to (47) the maximum value ρ_m of ρ, i.e. the value for closest approach of the ray to the sun, corresponds to $\sigma = 1$. The corresponding value of the angle ϕ is

$$\phi_m = \frac{\pi}{2}+\frac{\alpha}{\Delta}. \tag{51}$$

Thus, in the Euclidean picture mentioned above, the curve (50) is represented by a slightly curved line with a total deflexion

$$\psi = 2\left(\phi_m-\frac{\pi}{2}\right) = \frac{2\alpha}{\Delta}. \tag{52}$$

This will be the deflexion of a light ray in the field of the sun as noted by an observer on the earth, which in this connexion may be regarded as infinitely far from the sun. From (49) we get for the ρ-value corresponding to closest approach

$$\rho_m = \frac{1}{\Delta}\left(1+\frac{\alpha}{2\Delta}\right).$$

Hence, neglecting terms of second order in α, we get for (52)

$$\psi = 2\alpha\rho_m = \frac{2\alpha}{r_m} = \frac{4kM}{c^2 r_m}. \tag{53}$$

For a light ray which grazes the limb of the sun we then get, by means of (5′) and (6′), an angle of deflexion of 1·75″. This effect predicted by

Einstein† has been tested by observations during total eclipses of the sun on the apparent positions of stars whose light has passed close to the limb of the sun. The agreement between Einstein's formula (52) and the observations seems to be satisfactory,‡ but, as the effect is just inside the limits of experimental error, one cannot attach too much weight to the quantitative agreement.

The deflexion (52) is due partly to the varying velocity of light expressed by (40) and partly to the non-Euclidean character of the spatial geometry. If we had taken the space to be Euclidean, the equation (44) would have been replaced by

$$\dot{r}^2+r^2\dot{\phi}^2 = \frac{1}{c^2(1-\alpha/r)}. \tag{54}$$

Instead of (45) we would then have obtained

$$\left(\frac{d\rho}{d\phi}\right)^2 = \frac{1}{\Delta^2(1-\alpha\rho)}-\rho^2$$

which by integration to the first order in α leads to a deflexion

$$\psi = \alpha/\Delta \tag{55}$$

of only half of the value (52).

On the other hand, if the velocity of light is put equal to the constant c, while the spatial geometry is taken to be that given by the metric tensor $\gamma_{\iota\kappa}$ defined by (9) and (8), the trajectory of the light ray is, according to (38), a 'straightest' line. In this case the equations (44) and (45) are replaced by

$$\frac{\dot{r}^2}{1-\alpha/r}+r^2\dot{\phi}^2 = \frac{1}{c^2}, \qquad \left(\frac{d\rho}{d\phi}\right)^2 = (1-\alpha\rho)\left(\frac{1}{\Delta^2}-\rho^2\right)$$

respectively. Integration of the last equation then again leads to a deflexion

$$\psi = \alpha/\Delta \tag{56}$$

to the first order in α. Adding the two effects (55) and (56) we thus come back to Einstein's formula (52).

The direct experimental verifications of the general theory of relativity are thus very few; it should be remembered, however, that the general theory is not only a natural, but a nearly cogent generalization of the experimentally well-founded special theory. Further, since Einstein's gravitational theory contains Newton's theory as a first approximation, all the numerous astronomical observations which confirm the predictions of Newton's theory may therefore in a certain sense also

† A. Einstein, *Berl. Ber.*, p. 831 (1915); *Ann. d. Phys.* **49**, 769, § 22 (1916).
‡ W. W. Campbell and R. Trumpler, *Lick Observ. Bull.* **11**, 41 (1923); **13**, 130 (1928).

be regarded as a support of the general theory of relativity. The fact that the differences between the two theories become apparent only in the three small effects discussed above simply shows that Newton's theory represents an extremely good approximation for all gravitational phenomena inside the solar system. However, as regards cosmological problems which concern the structure and motion of larger parts of the universe, the two theories of gravitation can be expected to lead to very different results and, on account of its inner consistency and the generality of its basic principles, Einstein's gravitational theory may be expected to provide a safer guidance in the handling of these difficult problems.

132. Cosmological models

It has been known for a long time† that Newton's gravitational theory meets with serious difficulties when applied to the universe as a whole. Since Einstein's gravitational theory can be expected to give results which deviate appreciably from Newton's theory just for systems of cosmological extension it is clearly of great interest to investigate the new possibilities for a treatment of the universe as a whole, offered by the general theory of relativity. This question was taken up by Einstein‡ shortly after the development of the general theory and has since then been the object of many investigations by numerous authors. We shall not attempt here to give a detailed account of all these investigations, but confine ourselves to the consideration of the static homogeneous models of the universe originally proposed by Einstein§ and by de Sitter.‖

All models of the universe which have been considered so far are based on the assumption that the world, when looked at from a large-scale point of view, is spatially homogeneous and isotropic. It is true that the matter in the universe is partly gathered into stars which again have a tendency to cluster into nebulae of the same character as our own galaxy of stars. But, in the portion of space which can be reached by means of the present telescopes, these nebulae seem on the whole to be fairly uniformly distributed, and the assumption that the large-scale properties of the universe can be properly described by treating the matter as a perfect homogeneous fluid seems to be a natural starting-point. In the models proposed by Einstein and de Sitter the universe is

† C. Neumann, *Kgl sachs Ges d Wiss zu Leipzig, Math.-nat. Kl.* **26**, 97 (1874), H. v Seeliger, *Astron. Nachr.* **137**, p. 129 (1895), *Munch. Ber.* **26**, 373 (1896).

‡ A. Einstein, *Berl. Ber*, p. 142 (1917), *Ann. d. Phys* **55**, 241 (1918)

§ A Einstein, *Berl Ber.*, p 142 (1917).

‖ W. de Sitter, *Amst. Proc.* **19**, 1217 (1917); **20**, 229 (1917).

furthermore assumed to be a *static* system, which means that we can introduce a system of coordinates $x^i = (r, \theta, \phi, ct)$ in which the line element has the static and spherically symmetric form (XI. 63)

$$ds^2 = a(r)\,dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2) - b(r)c^2\,dt^2, \tag{57}$$

where a and b are functions of r only. On account of the assumed homogeneity of the universe any point in space may be taken as the origin $r = 0$ of the spatial system of coordinates

The functions $a(r)$ and $b(r)$ are now connected with the proper mass density $\overset{0}{\mu}{}^0$ and the proper pressure $\overset{0}{p}$ in the universe by the field equations (XI. 96, 94) for a perfect fluid,

$$\frac{d\overset{0}{p}}{dr} + (\overset{0}{\mu}{}^0c^2 + \overset{0}{p})\frac{b'}{2b} = 0, \tag{58}$$

$$\frac{b'}{abr} - \frac{1}{r^2}\left(1 - \frac{1}{a}\right) + \lambda = \kappa\overset{0}{p}, \tag{59 a}$$

$$\frac{a'}{a^2r} + \frac{1}{r^2}\left(1 - \frac{1}{a}\right) - \lambda = \kappa\overset{0}{\mu}{}^0c^2 \tag{59 b}$$

Here, $\overset{0}{p}$ and $\overset{0}{\mu}{}^0$ are constants on account of the assumed homogeneity of the model and we shall look for the possible regular solutions of these equations.

Since $d\overset{0}{p}/dr = 0$, (58) reduces to

$$(\overset{0}{\mu}{}^0c^2 + \overset{0}{p})b' = 0, \tag{60}$$

which requires either
$$b' = 0 \tag{61}$$
or
$$\overset{0}{\mu}{}^0c^2 + \overset{0}{p} = 0 \tag{62}$$

These two alternatives lead to the solutions of Einstein and de Sitter, respectively.†

133. The Einstein universe

The solution characterized by the condition (61) corresponds to a constant value of b, and by a suitable choice of the time variable t this constant may be made equal to 1, i e

$$b = 1. \tag{63}$$

Introducing (61) into (59 a) and solving this equation with respect to a, we obtain

$$a = \frac{1}{1-(\lambda-\kappa\overset{0}{p})r^2} = \frac{1}{1-r^2/R^2}, \tag{64}$$

† R. C Tolman, *Proc Nat Acad* **15**, 297 (1929)

where we have introduced a new constant R by the equation

$$\frac{1}{R^2} = \lambda - \kappa \overset{0}{p}. \tag{65}$$

Further, using (64) and (59 *b*), we get

$$\frac{1}{R^2} = \tfrac{1}{2}\kappa(\overset{0}{\mu}{}^0 c^2 + \overset{0}{p}), \tag{66}$$

which together with (65) leads to the relation

$$\lambda = \tfrac{1}{2}\kappa(\overset{0}{\mu}{}^0 c^2 + 3\overset{0}{p}). \tag{67}$$

Now the quantity $\overset{0}{\mu}{}^0$ is essentially positive and even if we allow for possible cohesive forces in the fluid giving rise to a negative value of $\overset{0}{p}$, the order of magnitude of $\overset{0}{p}$ will for any reasonable properties of the matter be far below the value $\overset{0}{\mu}{}^0 c^2$. Hence the constants λ and $1/R^2$ must also be positive and R will therefore be a real quantity of the dimension of a length.

In his original paper Einstein assumed the matter in the universe to be incoherent, thus exerting no pressure at all. In this case, we get from (66) and (67)

$$\lambda = \frac{1}{R^2} = \frac{\kappa c^2}{2}\overset{0}{\mu}{}^0. \tag{68}$$

On the other hand, if the universe is assumed to be mainly filled with radiation, we have, according to (VII. 145),

$$\overset{0}{p} = \frac{\overset{0}{\mu}{}^0 c^2}{3}, \tag{69}$$

which leads to the relation

$$\lambda = \frac{3}{2}\frac{1}{R^2} = \kappa \overset{0}{\mu}{}^0 c^2. \tag{70}$$

According to the estimate of Hubble,† the lower limit for the mean density of matter in the actual universe is of the order

$$\overset{0}{\mu}{}^0 \approx 10^{-30}\ \text{gm./cm.}^3$$

From (68) and (XI. 24) we then get

$$\lambda = \frac{1}{R^2} \approx 9 \times 10^{-58}\ \text{cm.}^{-2},$$

and, as an upper limit for R, we obtain

$$R \approx 3 \times 10^{28}\ \text{cm.} \approx 3 \times 10^{10}\ \text{light years} \tag{71}$$

† E Hubble, *Astrophys. Journ.* **79**, 8 (1934).

According to (57), (63), and (64) the line element of the Einstein universe is

$$ds^2 = \frac{dr^2}{1-r^2/R^2}+r^2(d\theta^2+\sin^2\theta\, d\phi^2)-c^2\, dt^2. \tag{72}$$

In a region of space for which

$$\frac{r^2}{R^2} \ll 1, \tag{73}$$

this line element reduces to that of the special theory of relativity. With the value of λ and R given by (71) we get for the distance $r \approx 10^{14}$ of the planet Neptune

$$\frac{r^2}{R^2} \approx \left(\frac{10^{14}}{10^{28}}\right)^2 = 10^{-28},$$

which means that the condition (73) is amply satisfied inside the solar system. This gives the justification for neglecting the λ-term in the treatment of all gravitational phenomena connected with the motion of the planets Furthermore, we see that the line element of the special theory of relativity, which is experimentally known to hold in every system of inertia far away from gravitating bodies, appears in the Einstein model as a consequence of the gravitational field equations for a homogeneous static distribution of the distant celestial bodies.

For larger parts of the universe, where the condition (73) is not satisfied, we get from (72) for the spatial metric tensor and the dynamical gravitational potentials

$$\left.\begin{aligned} \gamma_{11} &= \frac{1}{1-r^2/R^2}, \qquad \gamma_{22} = r^2, \qquad \gamma_{33} = r^2\sin^2\theta \\ \gamma_{\iota\kappa} &= 0 \text{ for } \iota \neq \kappa, \qquad \gamma = |\gamma_{\iota\kappa}| = \frac{r^4\sin^2\theta}{1-r^2/R^2} \end{aligned}\right\}, \tag{74}$$

$$\gamma_\iota = 0, \qquad \chi = 0. \tag{75}$$

The spatial line element

$$d\sigma^2 = \frac{dr^2}{1-r^2/R^2}+r^2\, d\theta^2+r^2\sin^2\theta\, d\phi^2 \tag{76}$$

is real only for $r < R$, which defines the extension of the physical space in the Einstein universe. The total volume of the universe is

$$V = \int_0^R\int_0^\pi\int_0^{2\pi} \sqrt{\gamma}\, drd\theta d\phi = \int_0^R\int_0^\pi\int_0^{2\pi} \frac{r^2\sin\theta\, drd\theta d\phi}{\sqrt{(1-r^2/R^2)}}$$

$$= 4\pi\int_0^R \frac{r^2\, dr}{\sqrt{(1-r^2/R^2)}} = \pi^2R^3 \tag{77}$$

and the greatest distance from the origin is obviously

$$L = \int_0^R \frac{dr}{\sqrt{(1-r^2/R^2)}} = R\int_0^1 \frac{d\xi}{\sqrt{(1-\xi^2)}} = R\sin^{-1}\xi\Big|_0^1 = \frac{\pi R}{2}. \tag{78}$$

Further, since the matter is at rest in our frame of reference, we get on account of (X. 56) for the total proper mass in the universe

$$M_0 = \int \overset{0}{\mu}{}^0\, d\overset{0}{V}{}^0 = \overset{0}{\mu}{}^0 \int d\overset{0}{V} = \pi^2 R^3 \overset{0}{\mu}{}^0, \tag{79}$$

and the mean distance of the matter from the origin is of the order

$$\frac{L}{2} \approx R.$$

By means of (79) and (68) or (70) we get

$$\lambda R^2 \approx 1, \qquad \frac{M_0 \kappa c^2}{4\pi R} \approx 1. \tag{80}$$

These relations are in accordance with the relations (XI. 49, 50) suggested by the considerations in § 120 on the nature of the centrifugal and Coriolis forces

The spatial geometry defined by (74)–(79) is of the so-called *elliptical* kind. However, the line element (76) allows also a somewhat different interpretation as regards the geometry of the universe as a whole. If we define four new variables y_0, y_1, y_2, y_3 by

$$\left.\begin{aligned} y_0 &= R\sqrt{(1-r^2R/^2)} \\ y_1 &= r\cos\theta \\ y_2 &= r\sin\theta\cos\phi \\ y_3 &= r\sin\theta\sin\phi \end{aligned}\right\}, \tag{81}$$

where
$$y_0^2+y_1^2+y_2^2+y_3^2 = R^2, \tag{82}$$

the line element (76) takes the Euclidean form

$$d\sigma^2 = dy_0^2+dy_1^2+dy_2^2+dy_3^2 \tag{83}$$

This shows that the physical space of the Einstein universe may also be interpreted as the three-dimensional surface of the sphere (82) of radius R in a four-dimensional Euclidean space with the Cartesian coordinates (y_0, y_1, y_2, y_3).

Defining polar coordinates R, ψ, θ, ϕ on the sphere (82) by

$$\left.\begin{aligned} y_0 &= R\cos\psi, \qquad & y_1 &= R\sin\psi\cos\theta \\ y_2 &= R\sin\psi\sin\theta\cos\phi, \qquad & y_3 &= R\sin\psi\sin\theta\sin\phi \end{aligned}\right\}, \tag{84}$$

we get
$$r = R\sin\psi \tag{85}$$

and to each point on the sphere corresponds one set of values of the variables (ψ, θ, ϕ) in the intervals

$$0 \leqslant \psi \leqslant \pi, \qquad 0 \leqslant \theta \leqslant \pi, \qquad 0 \leqslant \phi < 2\pi. \tag{86}$$

By using (84) in (76) the line element assumes the form

$$d\sigma^2 = R^2(d\psi^2 + \sin^2\psi\, d\theta^2 + \sin^2\psi \sin^2\theta\, d\phi^2), \tag{87}$$

and since in these coordinates

$$\gamma = |\gamma_{\iota\kappa}| = R^6 \sin^4\psi \sin^2\theta,$$

we get for the total volume of the universe

$$V = \int\limits_0^\pi d\psi \int\limits_0^\pi d\theta \int\limits_0^{2\pi} d\phi\, R^3 \sin^2\psi \sin\theta = 4\pi R^3 \int\limits_0^\pi \sin^2\psi\, d\psi = 2\pi^2 R^3. \tag{88}$$

Thus the volume in this so-called *spherical* space is twice the volume in the corresponding elliptical space. This is also clear from (85) and (86) which shows that the elliptical space covers only the hemisphere corresponding to $0 \leqslant \psi \leqslant \frac{1}{2}\pi$ or, in other words, antipodal points on the sphere are counted as one point only in elliptical space.

Similarly we find from (87) that the total distance around the closed spherical Einstein universe is

$$L = 2R \int\limits_0^\pi d\psi = 2\pi R, \tag{89}$$

which thus represents the distance one would have to travel along a 'major circle' on the sphere (82) in order to return to the starting-point.

The relations (80) remain true also in the spherical Einstein universe.

Since the dynamical potentials (γ_ι, χ) are zero, the gravitational force on a particle is zero and since, further, the system of reference is rigid, *a free particle in the universe is moving with constant velocity along the straightest line compatible with the spatial geometry of the universe* (see end of § 110) In this generalized form the law of inertia is thus valid also in the Einstein universe. A free particle at rest will therefore remain at rest.

According to (X. 94) and (75) the velocity of light is constant and equal to c. Light emitted by a star at rest at the point (r, θ, ϕ), at the time t_1 will therefore arrive at the origin $r = 0$ at the time

$$t_2 = t_1 + \frac{1}{c} \int\limits_0^{r_1} \frac{dr}{\sqrt{(1 - r^2/R^2)}} = t_1 + \frac{R}{c} \sin^{-1}\frac{r_1}{R}, \tag{90}$$

and since r_1 is constant, we get by differentiation

$$\Delta t_2 = \Delta t_1. \tag{91}$$

Thus the time-interval Δt_2 between the arrival of two successive wave crests at the origin is equal to the interval Δt_1 between their emission. Further, since $\chi = 0$, the time variable t is identical with the time shown by standard clocks at rest and (91) therefore means that the frequency of the light as determined by an observer at the origin is equal to the proper frequency of the light emitted by a star at rest.

Apart from small Doppler effects due to individual random motions of the distant nebulae we should therefore not expect any systematic shift of the spectral lines emitted by the nebulae. In the actual universe, however, the work of Hubble and Humason† shows a definite red shift in the light from nebulae, which increases linearly with the distance. This clearly shows that the Einstein model, in spite of its many satisfactory features, represents only an approximate description of the actual universe.

134. The de Sitter universe

Besides the solution corresponding to the condition (61) which leads to the Einstein model there exists another static homogeneous and isotropic solution of the general field equations arising from the condition (62), viz.

$$\mathring{\mu}^0 c^2 + \mathring{p} = 0. \tag{92}$$

If we add the equations (59 a) and (59 b) we get in this case

$$(ab)' = \kappa(\mathring{\mu}^0 c^2 + \mathring{p})a^2 br = 0$$

or

$$ab = \text{constant}.$$

By a trivial change of scale of the time variable this constant can of course always be made equal to 1, which means that b is the reciprocal of a:

$$ab = 1. \tag{93}$$

Introducing $y = 1/a$ as a new variable, (59 b) can now be written

$$(yr)' \equiv y'r + y = +1 - (\lambda + \kappa\mathring{\mu}^0 c^2)r^2,$$

which by integration gives

$$yr = r - \frac{(\lambda + \kappa\mathring{\mu}^0 c^2)}{3} r^3 + \text{constant}. \tag{94}$$

† E. Hubble, *Proc. Nat. Acad.* **15**, 168 (1929), E. Hubble and M. L. Humason, *Astrophys. Journ.* **74**, 43 (1931).

Since y is regular for $r = 0$ the constant on the right-hand side of (94) must be zero. Thus we get from (93) and (94)

$$b = \frac{1}{a} = y = 1 - \frac{r^2}{R^2}, \tag{95}$$

where we have put
$$\frac{1}{R^2} = \frac{\lambda + \kappa \mathring{\mu}^0 c^2}{3}. \tag{96}$$

Using (95) in (57) we then get the de Sitter line element in the form

$$ds^2 = \frac{dr^2}{1 - r^2/R^2} + r^2(d\theta^2 + \sin^2\theta \, d\phi^2) - \left(1 - \frac{r^2}{R^2}\right) c^2 \, dt^2. \tag{97}$$

The components of the spatial metric tensor are again given by (74), i.e. the spatial geometry in the de Sitter universe is of the same form as in the Einstein model, at least if λ is assumed to be positive. However, the dynamical potentials are not zero in this case; instead we have

$$\gamma_\iota = 0, \qquad \chi = -\frac{r^2 c^2}{2R^2}. \tag{98}$$

Thus a free particle is acted upon by a gravitational force

$$\mathbf{K} = -m \operatorname{grad} \chi = m\left(\frac{rc^2}{R^2}, 0, 0\right) \tag{99}$$

proportional to the variable r, which means that the law of inertia does not hold over large regions of space in the de Sitter world. Only in regions for which

$$\frac{r^2}{R^2} \ll 1$$

will the line element (97) again reduce to the line element of the special theory of relativity and the law of inertia is approximately valid.

The equations of motion of a free particle may now be obtained from (X. 15) by using the expression for the gravitational force given by (99). It is, however, more convenient to make use of the possibility discussed in § 113 of eliminating the dynamical potentials by the introduction of a suitable set of space-time coordinates. This is attained by defining new variables r', θ', ϕ', t' by the transformation

$$\left.\begin{aligned} r' = \frac{r}{\sqrt{(1 - r^2/R^2)}} e^{-ct/R}, \qquad t' = t + \frac{R}{2c} \ln\left(1 - \frac{r^2}{R^2}\right) \\ \theta' = \theta, \qquad \phi' = \phi \end{aligned}\right\}. \tag{100}$$

As shown independently by Lemaître† and by Robertson,‡ this transformation leads to the following form for the line element:

$$ds^2 = e^{2ct'/R}(dr'^2+r'^2\,d\theta'^2+r'^2\sin^2\theta'\,d\phi'^2)-c^2\,dt'^2, \tag{101}$$

which is easily verified by direct calculation. Finally, defining new space-coordinates x', y', z' connected with r', θ', ϕ' by the usual equations connecting Cartesian coordinates and polar coordinates in a Euclidean space, (101) may be written

$$ds^2 = e^{2ct'/R}(dx'^2+dy'^2+dz'^2)-c^2\,dt'^2. \tag{102}$$

The new coordinates x', y', z', t' can take all values from $-\infty$ to $+\infty$.

With these coordinates we have

$$\left.\begin{aligned} \gamma_{11} = \gamma_{22} = \gamma_{33} = \frac{1}{\gamma^{11}} = \frac{1}{\gamma^{22}} = \frac{1}{\gamma^{33}} = e^{2ct'/R} \\ \gamma_{\iota\kappa} = \gamma^{\iota\kappa} = 0 \text{ for } \iota \neq \kappa \\ \gamma_{\iota} = 0, \qquad \chi = 0 \end{aligned}\right\}, \tag{103}$$

and the spatial Christoffel symbols are all zero.

Thus, the time variable t' is the time shown by a standard clock at rest at any reference point. At any fixed time t' the spatial geometry is Euclidean, x', y', z' being Cartesian coordinates apart from the common factor $e^{2ct'/R}$. The distance from the origin $r' = 0$ to a point (r',θ',ϕ'), as measured by standard measuring-rods, is

$$l = e^{ct'/R}r'. \tag{104}$$

The velocity of light is constant and equal to c.

$$w = \frac{d\sigma}{dt'} = c \tag{105}$$

Further, it is easily seen that the trajectories of light rays are straight lines. This follows at once from the first equations (VIII 87) for the time track of a light ray. Taking $i = \iota = 1, 2, 3$ in these equations, we get

$$\frac{d}{d\lambda}\left(e^{2ct'/R}\frac{dx'^{\iota}}{d\lambda}\right) = 0,$$

and by integration

$$\frac{dx'^{\iota}}{d\lambda} = a^{\iota}e^{-2ct'/R}, \tag{106}$$

where the a^{ι} are constants of integration. Hence

$$\frac{dx'/dt'}{a^1} = \frac{dy'/dt'}{a^2} = \frac{dz'/dt'}{a^3} \tag{107}$$

† G. E. Lemaître, *J. Math. and Phys.* (*M.I.T.*), **4**, 188 (1925).
‡ H. P. Robertson, *Phil. Mag.* **5**, 835 (1928).

and the equations of the trajectory of a light ray are linear equations of the form

$$\frac{x'-x'_0}{a^1} = \frac{y'-y'_0}{a^2} = \frac{z'-z'_0}{a^3}. \tag{108}$$

Since the light rays are rectilinear in the (x', y', z')-space we can obviously apply the usual triangulation method to determine the parallax of celestial bodies and in this way determine the distance (104) by direct measurements.

Let us now consider the propagation of a light ray along the x'-axis. From (105) and (103) we get

$$e^{ct'/R}\frac{dx'}{dt'} = \pm c, \tag{109}$$

or

$$\frac{dx'}{dt'} = \pm ce^{-ct'/R}. \tag{110}$$

The motion of a light signal starting off at the point $x' = x'_0 > 0$ at the time t'_0 in the direction of the negative x'-axis is thus given by

$$x' = x'_0 + R(e^{-ct'/R} - e^{-ct'_0/R}), \tag{111}$$

and we see that, unless

$$x'_0 < Re^{-ct'_0/R}, \tag{112}$$

this signal will never reach the origin $x' = 0$.

Although the world corresponding to the line element (102) is apparently infinite, an observer at the origin will never be able to obtain any information about the regions outside the 'horizon' defined by (112). The greatest distance inside the observable world of an observer placed at the origin at the time t'_0 is thus, by (104) and (112),

$$L = e^{ct'_0/R}Re^{-ct'_0/R} = R, \tag{113}$$

i.e. it is the same for all times in accordance with the intrinsically static character of the de Sitter universe. Further, since (102) is invariant against any displacement of the origin, this consideration holds for any observer at rest in the system of reference considered here. However, the position of the horizon will of course be different for the different 'equivalent' observers.

In the present coordinates the gravitational force on a free particle is zero and the equations of motion are given by (X. 20′):

$$\frac{d_c p_i}{dt'} = 0, \tag{114}$$

where

$$\left.\begin{aligned} p_\iota &= mu_\iota \\ m &= \frac{\overset{0}{m}_0}{\sqrt{(1-u^2/c^2)}} \\ u_\iota &= \gamma_{\iota\kappa}\frac{dx'^\kappa}{dt'} \end{aligned}\right\}. \tag{115}$$

Further, since the spatial Christoffel symbols are zero in the present case, we have

$$\frac{d_c p_\iota}{dt'} = \frac{dp_\iota}{dt'} = 0,$$

which shows that the covariant components of the momentum vector are constant:

$$p_\iota = \text{constant}. \tag{116}$$

However, the contravariant components are not constant, since we have

$$p^\iota = \gamma^{\iota\kappa}p_\kappa = e^{-2ct'/R}p_\iota \tag{117}$$

in accordance with the differential equations (X. 19′). From (117) we now get

$$\left.\begin{aligned} mu^\iota = m\frac{dx'^\iota}{dt'} = p_\iota e^{-2ct'/R} \\ \text{or} \qquad \frac{dx'/dt'}{b^1} = \frac{dy'/dt'}{b^2} = \frac{dz'/dt'}{b^3} \end{aligned}\right\}, \tag{118}$$

where b^1, b^2, b^3 are constants. Thus the orbits of free particles are also straight lines described by equations of the form

$$\frac{x'-x'_0}{b^1} = \frac{y'-y'_0}{b^2} = \frac{z'-z'_0}{b^3}, \tag{119}$$

but the velocity of the particle is in general not constant. Only if the velocity at one time is zero relative to the system of reference considered here will it remain zero; for from (118) we see that

$$\frac{dx'^\iota}{dt'} = 0 \quad \text{for } t' = t'_0$$

requires $p_\iota = 0$ and dx'^ι/dt' will then be zero for all times.

Any reference point $(x', y', z') =$ constant may thus be represented by a freely falling particle in accordance with the remarks at the end of § 113. Since the system of reference is not rigid, the distance from the origin of a point with constant values of the spatial coordinates will, however, depend on the time according to (104). Hence, if we assume that the nebulae may be treated as free particles at rest in the present system of coordinates, their distance from the origin increases with a radial velocity

$$v_r = \frac{dl}{dt'} = \frac{c}{R} e^{ct'/R} r' = \frac{c}{R} l, \tag{120}$$

which is proportional to the distance l.

This radial velocity of the nebulae must be expected to give rise to a shift towards the red of the spectral lines emitted by the nebulae. To investigate this question we consider light emitted by a particle which is permanently located at the reference point (r', θ', ϕ'). According to (105) and (103) the radial velocity of light is

$$\frac{dr'}{dt'} = \pm c e^{-ct'/R}. \tag{121}$$

Thus, if t_1' and t_2' are the times for the emission of radiation from the particle and its reception at the origin $r' = 0$, respectively, we get by integration

$$\int_{r'}^{0} dr' = -c \int_{t_1'}^{t_2} e^{-ct'/R}\, dt' \quad \text{or} \quad r' = R(e^{-ct_1'/R} - e^{-ct_2'/R}). \tag{122}$$

On differentiation we get for the time-interval Δt_2 between the arrival of two successive wave crests at the origin and the time-interval Δt_1 between their emission the following relation:

$$\Delta t_2' = \Delta t_1'\, e^{c(t_2 - t_1)/R}. \tag{123}$$

Now, since t' is the time shown by a standard clock at rest and since the velocity of light is c everywhere, we have for the proper wave-length λ^0 of the emitted light

$$\lambda^0 = c\, \Delta t_1'.$$

Similarly the wave-length as measured by an observer at the origin will be

$$\lambda = c\, \Delta t_2'.$$

Hence we get from (123)

$$\lambda = \lambda^0 e^{c(t_2' - t_1')/R}. \tag{124}$$

Further, the distance to the particle at the time of observation of the light is, according to (104) and (122),

$$l = r' e^{ct_2'/R} = R(e^{c(t_2' - t_1')/R} - 1). \tag{125}$$

From (124) we therefore get

$$\frac{\Delta\lambda}{\lambda^0} = \frac{\lambda - \lambda^0}{\lambda^0} = \frac{l}{R}. \tag{126}$$

Thus, if we assume that the nebulae, apart from small individual velocities, are at rest in the system of coordinates used here, we get an explanation of the actual red shift observed by Huble and Humason To obtain quantitative agreement the constant R in (126) must have the value

$$R \simeq 1\cdot 66 \times 10^{27} \text{ cm.} \simeq 1\cdot 75 \times 10^{9} \text{ light years.} \tag{127}$$

This value for the 'radius' of the universe is somewhat smaller than the value (71) in the case of the Einstein universe. It is true that this explanation rests essentially on the assumption that the nebulae in the mean are at rest relative to the system of reference defined by the coordinates (x', y', z', t'). This assumption, known as Weyl's hypothesis,† is, however, very natural and has the attractive feature that all nebulae in this model are on the same footing so that an observer at any other reference point would observe the same red shift of the light coming from the nebulae as the observer at the arbitrarily chosen origin of our system of coordinates.

In the original system of coordinates the motion of the nebulae is obtained from (100) by solving the first equation with respect to r. Hence

$$r^2 = \frac{r'^2}{e^{-2ct/R}+r'^2/R^2}, \tag{128}$$

where r' is a constant for each nebula. Thus we see that the nebulae according to the Weyl hypothesis are freely falling particles which start at the point $r = 0$ at $t = -\infty$ and end at the antipodal point $r = R$ for $t \to +\infty$.

Finally introducing five variables $z_\mu = (z_0, z_1, z_2, z_3, z_4)$ by the equations

$$\left.\begin{aligned} z_1 = x'e^{ct'/R}, \qquad z_2 = y'e^{ct'/R}, \qquad z_3 &= z'e^{ct'/R} \\ z_0 = R\left(\cosh\frac{ct'}{R} - \frac{r'^2}{2R^2}e^{ct'/R}\right) \\ z_4 = iR\left(\sinh\frac{ct'}{R} + \frac{r'^2}{2R^2}e^{ct'/R}\right) \end{aligned}\right\}, \tag{129}$$

we have

$$\sum_{\mu=0}^{4} z_\mu^2 = R^2, \tag{130}$$

and the line element takes the form

$$ds^2 = \sum_{\mu=0}^{4} (dz_\mu)^2. \tag{131}$$

The space-time continuum of the de Sitter universe may thus be pictured as the four-dimensional surface of a sphere of radius R in a five-dimensional pseudo-Euclidean space just as the 4-space of the Einstein

† H. Weyl, *Phys. ZS.* **24**, 230 (1923), *Phil Mag.* **9**, 936 (1930).

universe, according to (72), (76), and (83), can be pictured as a cylinder with spherical cross-section and the axis in the direction of the time-axis.

Since the equations (130) and (131) are form-invariant under the group of five-dimensional orthogonal transformations of the variables (z_μ), the line element (102) will be form-invariant under the group L of the corresponding transformations of the variables $x^i = (x', y', z', ct')$ connected with $(z_0,..., z_4)$ by the equations (129). The transformations of the group L thus connect systems of coordinates (x^i) which are equivalent in the sense of § 121. These transformations play the same role in the de Sitter universe as the inhomogeneous Lorentz transformations in the flat space of the special theory of relativity.†

Although the de Sitter model leads to a natural explanation of the observed red shift of spectral lines, this model can hardly be regarded as a satisfactory model of the actual universe for the following reason. According to the condition (92) underlying this model, the density and pressure of the celestial matter which give rise to the line element (97) satisfy the equation

$$\mathring{\mu}^0 c^2 + \mathring{p} = 0. \tag{132}$$

Since $\mathring{\mu}^0 c^2 \gg 0$, (132) leads to the assumption that $\mathring{p}$ is negative and very large. Even if we permit the existence of cohesive forces in the ideal fluid filling our model, a value of $\mathring{p}$ of the order of $-\mathring{\mu}^0 c^2$ would, however, be quite incompatible with the properties of any known material. Hence (132) can be satisfied only if we take the density to be zero or at least very much smaller than the mean density of the actual celestial matter. Hence the de Sitter model corresponds to an empty universe containing no appreciable amount of matter and radiation, the stars and nebulae being treated then as a kind of test bodies which do not contribute essentially to the gravitational field. This point of view is, however, not in agreement with the basic ideas underlying the general theory of relativity according to which the centrifugal forces and Coriolis forces, for instance, are due to the motion of the distant celestial bodies relative to the rotating system. While the non-permanent gravitational fields can be explained along these lines in the Einstein model of the universe, the empty de Sitter model does not of course afford any basis for such explanation, the non-permanent fields being here of the same nature as the fictitious forces in Newton's theory.

These considerations make it probable that the Einstein model represents a better approximation to the actual universe than the de Sitter

† See ref., Chap. XII, p. 364, and C. Møller, *Dan. Mat. Fys. Medd.* 18, No. 6, p. 40 (1941).

model despite its failure to explain the systematic red shift of the light coming from the nebulae. If we want to construct a model in which the advantages of the two static models of Einstein and de Sitter are combined we must obviously have recourse to non-static models in which the metric tensor is intrinsically time-dependent. Such models have been extensively studied in the literature.† These investigations include models which expand from an originally static state as well as models which undergo successive expansions and contractions. Our present knowledge of the actual universe, which only covers a limited region in space and time, is, however, totally insufficient and thus no unique choice between the different non-static models is possible.

† See the extensive treatment of these problems in Tolman's book, *Relativity, Thermodynamics, and Cosmology*, Oxford, 1934

APPENDIXES

1. Gauss's theorem

LET V be a finite domain bounded by a closed surface σ in a three-dimensional Euclidean space and let $f(x_1, x_2, x_3)$ be a given function of the Cartesian coordinates x_1, x_2, x_3 inside V. Consider the volume integral $\int_V \frac{\partial f}{\partial x_1}\, dx_1\, dx_2\, dx_3$ over the region V. Assuming that the surface σ is convex so that a straight line parallel to the x_1-axis with constant values of x_2 and x_3 intersects the surface σ in two points only, this integral may, by partial integration with respect to the variable x_1, be written

$$\int_V \frac{\partial f}{\partial x_1}\, dx_1\, dx_2\, dx_3 = \int_{\sigma_1} (f^+ - f^-)\, dx_2\, dx_3, \tag{1}$$

where the integration on the right-hand side is extended over the projection σ_1 of the surface σ on the $(x_2 x_3)$-plane, and f^+ and f^- are the values of f at the two points p^+ and p^- on σ whose projections on the $(x_2 x_3)$-plane lie inside the face element $dx_2\, dx_3$ of σ_1. Now let $d\sigma^+$, $d\sigma^-$ be the surface elements of σ at p^+ and p^-, respectively, whose projections on the $(x_2 x_3)$-plane are equal to the element $dx_2\, dx_3$. Then, if $\mathbf{n}^+ = (n_1^+, n_2^+, n_3^+)$ and $\mathbf{n}^- = (n_1^-, n_2^-, n_3^-)$ are unit vectors in the direction of the outward normals of $d\sigma^+$ and $d\sigma^-$ and if $x_1^+ > x_1^-$, we obviously have

$$n_1^+\, d\sigma^+ = -n_1^-\, d\sigma^- = dx_2\, dx_3 \tag{2}$$

Hence (1) may be written

$$\int_V \frac{\partial f}{\partial x_1}\, dV = \int_\sigma f . n_1\, d\sigma, \tag{3}$$

where the integration on the right-hand side is extended over the whole surface σ, and n_1 is the x_1-component of the *outward* normal vector $\mathbf{n}$. It is easily seen that (3) holds also in the case where the boundary surface σ is not convex, in which case a straight line parallel to the x_1-axis may intersect the surface in a greater—but *even*—number of points. Similarly, if $g(x_1, x_2, x_3)$, $h(x_1, x_2, x_3)$ are two other given functions of the space coordinates we also have

$$\left.\begin{aligned} \int_V \frac{\partial g}{\partial x_2}\, dV &= \int_\sigma g . n_2\, d\sigma, \\ \int_V \frac{\partial h}{\partial x_3}\, dV &= \int_\sigma h . n_3\, d\sigma. \end{aligned}\right\} \tag{3'}$$

If, in particular, f, g, and h are the components $a_1(\mathbf{x})$, $a_2(\mathbf{x})$, $a_3(\mathbf{x})$ of a vector field $\mathbf{a} = (a_i)$ we get by addition of the three equations (3), (3')

$$\int_V \frac{\partial a_i}{\partial x_i}\, dV = \int_\sigma (a_i n_i)\, d\sigma$$

or

$$\int_V \operatorname{div} \mathbf{a}\, dV = \int_\sigma (\mathbf{a} \cdot \mathbf{n})\, d\sigma = \int_\sigma a_n\, d\sigma, \tag{4}$$

where a_n is the component of $\mathbf{a}$ in the direction of the outward normal. The equation (4) which enables us to transform the volume integral on the left-hand side into a surface integral represents Gauss's theorem (IV. 188).

Further, if $t_{\iota\kappa}(\mathbf{x})$ is a three-dimensional tensor field of rank 2, we get, putting $t_{\iota 1} = f$, $t_{\iota 2} = g$, $t_{\iota 3} = h$ from (3) and (3′)

$$\int\limits_V \frac{\partial t_{\iota\kappa}}{\partial x_\kappa} dV = \int\limits_\sigma (t_{\iota\kappa} n_\kappa)\, d\sigma, \tag{5}$$

i.e. the equation (IV. 192).

2. The transformation equations for the four-current density

According to (V. 4, 4′) the transformation equations connecting the four-current densities s_i and s_i' in two systems of inertia S and S' must be such that the equation

$$\frac{\partial s_i'}{\partial x_i'} = 0 \tag{1}$$

is a consequence of the equation

$$\frac{\partial s_i}{\partial x_i} = 0 \tag{1′}$$

for all possible charge and current distributions. The transformation must therefore be of the form

$$s_i' = f_i(s_1, s_2, s_3, s_4), \tag{2}$$

where the functions f_i must satisfy the relations

$$\frac{\partial f_i}{\partial x_i'} = \frac{\partial f_i}{\partial s_k}\frac{\partial s_k}{\partial x_l}\frac{\partial x_l}{\partial x_i'} = \left(\frac{\partial f_i(s_m)}{\partial s_k}\alpha_{il}\right)s_{k,l} = 0 \tag{3}$$

for all variations of the twenty variables s_i and $s_{k,l} = \partial s_k/\partial x_l$ which are subjected to the restriction

$$s_{i,i} = 0 \tag{4}$$

Multiplying (4) by a Lagrangian factor λ, which may be a function of the variables (s_i), and subtracting this equation from (3), we get the equation

$$\left(\frac{\partial f_i}{\partial s_k}\alpha_{il} - \lambda\,\delta_{k,l}\right)s_{k,l} = 0, \tag{5}$$

which must now hold for arbitrary independent variations of the variables s_i and $s_{k,l}$. By variation of the variables $s_{k,l}$ we get

$$\frac{\partial f_i}{\partial s_k}\alpha_{il} = \lambda(s_r)\delta_{kl}, \tag{6}$$

or, using the orthogonality relations (IV. 14),

$$\frac{\partial f_i}{\partial s_k} = \lambda(s_r)\alpha_{ik} \tag{6′}$$

From (6) we get by differentiation with respect to s_m

$$\frac{\partial^2 f_i}{\partial s_k\,\partial s_m}\alpha_{il} = \frac{\partial\lambda}{\partial s_m}\delta_{kl},$$

and since the left-hand side is symmetrical in k and m we must have

$$\frac{\partial\lambda}{\partial s_m}\delta_{kl} = \frac{\partial\lambda}{\partial s_k}\delta_{ml}.$$

Taking $m = l \neq k$ in this equation we get

$$\frac{\partial\lambda}{\partial s_k} = 0, \tag{7}$$

i.e. λ must be independent of the variables s_i also. Hence we get by integration of (6′)

$$f_i(s_k) = \lambda\alpha_{ik}s_k + \beta_{ik},$$

where the β_{ik} and λ are constants. If the electric charge density is zero in S, i.e. $s_i = 0$, it must also be zero in S'. Hence $\beta_{ik} = 0$ and the transformations (2) take the form

$$s'_i = \lambda\alpha_{ik}s_k. \tag{8}$$

From (IV. 11), (V. 3), and (8) we now get

$$s'_i s'_i = \lambda^2 s_k s_k,$$

or

$$\rho'^2(1-u'^2/c^2) = \lambda^2\rho^2(1-u^2/c^2). \tag{9}$$

By the same relativity argument as was used in connexion with the equations (II. 8, 14) it now follows that the constant λ must have the value

$$\lambda = 1, \tag{10}$$

and (8) then leads to the transformation (V 5), showing that the four-current density is a four-vector.

3. Plane waves in a homogeneous isotropic substance

In the rest system of a homogeneous isotropic substance with electric and magnetic constants ϵ and μ and $\rho = \mathbf{J} = 0$ Maxwell's equations (VII. 31, 32) may be written

$$\operatorname{div}\mathbf{H} = \operatorname{div}\mathbf{E} = 0, \tag{1}$$

$$\epsilon\dot{\mathbf{E}} = c\operatorname{curl}\mathbf{H}, \qquad \mu\dot{\mathbf{H}} = -c\operatorname{curl}\mathbf{E}, \tag{2}$$

$$\mathbf{D} = \epsilon\mathbf{E}, \qquad \mathbf{B} = \mu\mathbf{H}. \tag{3}$$

In the case of a plane wave with wave planes perpendicular to the x-axis of a Cartesian system of coordinates (x, y, z) the field vectors are functions of x and t only. From (1) and from the x-component of the vector equation (2) we therefore get

$$\frac{\partial E_x}{\partial x} = \frac{\partial H_x}{\partial x} = \dot{E}_x = \dot{H}_x = 0,$$

and since constant fields are of no importance in optics we may put

$$E_x = H_x = 0 \tag{4}$$

The y- and z-components of the equation (2) give

$$\epsilon\frac{\partial E_y}{\partial t} = -c\frac{\partial H_z}{\partial x}, \qquad \epsilon\frac{\partial E_z}{\partial t} = c\frac{\partial H_y}{\partial x}, \tag{5}$$

$$\mu\frac{\partial H_y}{\partial t} = c\frac{\partial E_z}{\partial x}, \qquad \mu\frac{\partial H_z}{\partial t} = -c\frac{\partial E_y}{\partial x}. \tag{6}$$

Differentiating one of the sets of equations (5) or (6) with respect to t and using the other set we see that each of the functions E_y, E_z, H_y, H_z satisfies the wave equation

$$\frac{\partial^2\psi}{\partial t^2} = w^2\frac{\partial^2\psi}{\partial x^2}, \tag{7}$$

where

$$w = \frac{c}{\sqrt{(\epsilon\mu)}} \tag{8}$$

is a constant

The general solution of the wave equation (7) is

$$\psi = f_1(t-x/w)+f_2(t+x/w),$$

where f_1 and f_2 are arbitrary functions. The two terms in this expression represent plane waves travelling with velocity w in the directions of the positive x-axis and negative x-axis, respectively. Thus considering a wave travelling in the direction of the positive x-axis, we can put

$$E_y = \epsilon^{-\frac{1}{2}} f(t - x/w), \qquad E_z = \epsilon^{-\frac{1}{2}} g(t - x/w), \tag{9}$$

where f and g are arbitrary functions of the argument $t - x/w$. From (6) we then get by integration

$$H_y = -\mu^{-\frac{1}{2}} g(t - x/w), \qquad H_z = \mu^{-\frac{1}{2}} f(t - x/w). \tag{10}$$

The equations (4), (9), and (10) represent the most general expression for a plane wave moving in the direction of the positive x-axis. Introducing three unit vectors

$$\mathbf{n} = (1, 0, 0),$$
$$\mathbf{e}^{(1)} = (0, 1, 0),$$
$$\mathbf{e}^{(2)} = (0, 0, 1),$$

these equations may be written

$$\left.\begin{aligned} \mathbf{E} &= \epsilon^{-\frac{1}{2}} f(t - (\mathbf{x}\cdot\mathbf{n})/w)\mathbf{e}^{(1)} + \epsilon^{-\frac{1}{2}} g(t - (\mathbf{x}\cdot\mathbf{n})/w)\mathbf{e}^{(2)}, \\ \mathbf{H} &= -\mu^{-\frac{1}{2}} g(t - (\mathbf{x}\cdot\mathbf{n})/w)\mathbf{e}^{(1)} + \mu^{-\frac{1}{2}} f(t - (\mathbf{x}\cdot\mathbf{n})/w)\mathbf{e}^{(2)}, \end{aligned}\right\} \tag{11}$$

and in this vector form they remain true if the system of coordinates is rotated so that $\mathbf{n}$ does not lie in the direction of the new x-axis. $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ are then arbitrary but constant unit vectors perpendicular to each other and to the wave normal $\mathbf{n}$.

4. Transformation of the gravitational field variables $\gamma_{\iota\kappa}, \gamma_\iota, \chi, \omega_{\iota\kappa}$ by a change of coordinates inside a definite system of reference

The transformation (VIII. 59)

$$x'^\iota = x'^\iota(x^\kappa), \qquad x'^4 = f(x^i), \tag{1}$$

connecting two different coordinate systems inside the same system of reference, is characterized by the equations

$$\alpha^\iota_4 = \frac{\partial x'^\iota}{\partial x^4} = 0, \qquad \breve{\alpha}^\iota_4 = \frac{\partial x^\iota}{\partial x'^4} = 0. \tag{2}$$

Thus we get from (VIII. 57)

$$g'_{\iota 4} = \breve{\alpha}^l_\iota \breve{\alpha}^m_4 g_{lm} = \breve{\alpha}^l_\iota \breve{\alpha}^4_4 g_{l4} = \breve{\alpha}^4_4(\breve{\alpha}^\lambda_\iota g_{\lambda 4} + \breve{\alpha}^4_\iota g_{44}), \tag{2'}$$

from which we obtain

$$\left.\begin{aligned} 1 + 2\chi'/c^2 &= -g'_{44} = (\breve{\alpha}^4_4)^2(-g_{44}) = (\breve{\alpha}^4_4)^2(1 + 2\chi/c^2) \\ \gamma'_\iota &= \frac{g'_{\iota 4}}{\sqrt{(-g'_{44})}} = \frac{\breve{\alpha}^4_4}{|\breve{\alpha}^4_4|}(\breve{\alpha}^\kappa_\iota \gamma_\kappa - \breve{\alpha}^4_\iota \sqrt{(-g_{44})}) \end{aligned}\right\}. \tag{3}$$

Further,

$$\gamma'_{\iota\kappa} = g'_{\iota\kappa} + \gamma'_\iota \gamma'_\kappa = \breve{\alpha}^l_\iota \breve{\alpha}^m_\kappa g_{lm} + (\breve{\alpha}^\lambda_\iota \gamma_\lambda - \breve{\alpha}^4_\iota \sqrt{(-g_{44})})(\breve{\alpha}^\mu_\kappa \gamma_\mu - \breve{\alpha}^4_\kappa \sqrt{(-g_{44})}),$$

i.e.

$$\gamma'_{\iota\kappa} = \breve{\alpha}^\lambda_\iota \breve{\alpha}^\mu_\kappa \gamma_{\lambda\mu}. \tag{4}$$

Thus $\gamma_{\iota\kappa}$ transforms like a spatial tensor and $d\sigma^2 = \gamma_{\iota\kappa}\,dx^\iota dx^\kappa$ is invariant under the transformation (1).

The transformation (VIII. 103)

$$x'^\iota = x^\iota, \qquad x'^4 = f(x^i) \tag{5}$$

is a special transformation (1) in which only the rate and setting of the coordinate clocks are changed. In this case we have

$$\alpha^\iota_k = \check{\alpha}^\iota_k = \frac{\partial x'^\iota}{\partial x^k} = \frac{\partial x^\iota}{\partial x'^k} = \delta^\iota_k. \tag{6}$$

Using this in (VIII. 55) with $i = 4$, we get

$$\alpha^4_\kappa + \alpha^4_4 \check{\alpha}^4_\kappa = 0, \qquad \alpha^4_4 \check{\alpha}^4_4 = 1, \tag{7}$$

i.e.

$$\frac{\partial f}{\partial x^\kappa} + \frac{\partial f}{\partial x^4}\check{\alpha}^4_\kappa = 0, \qquad \frac{\partial f}{\partial x^4}\check{\alpha}^4_4 = 1. \tag{8}$$

If we want the gravitational vector potentials γ'_ι to be zero in the primed system we get from (3) the conditions

$$\check{\alpha}^\kappa_\iota \gamma_\kappa - \check{\alpha}^4_\iota \sqrt{(1+2\chi/c^2)} = \gamma_\iota - \check{\alpha}^4_\iota \sqrt{(1+2\chi/c^2)} = 0, \tag{9}$$

which by means of (7) lead to the equations (VIII. 109)

$$\frac{\partial f}{\partial x^\iota} + \frac{\gamma_\iota}{\sqrt{(1+2\chi/c^2)}}\frac{\partial f}{\partial x^4} = 0. \tag{10}$$

Putting

$$\sigma_\iota = \frac{\gamma_\iota}{\sqrt{(-g_{44})}} = \frac{g_{\iota 4}}{-g_{44}}, \tag{11}$$

we get from (2), (2′), (6), and (8)

$$\sigma'_\iota = \frac{g'_{\iota 4}}{-g'_{44}} = \frac{\sigma_\iota}{\check{\alpha}^4_4} - \frac{\check{\alpha}^4_\iota}{\check{\alpha}^4_4} = \frac{\partial f}{\partial x^4}\sigma_\iota + \frac{\partial f}{\partial x^\iota}. \tag{12}$$

It is now easily seen that the operator $\frac{\partial}{\partial x^\iota} + \sigma_\iota \frac{\partial}{\partial x^4}$ is invariant under the transformation (5), for we have

$$\frac{\partial}{\partial x'^\iota} = \frac{\partial x^k}{\partial x'^\iota}\frac{\partial}{\partial x^k} = \delta^\kappa_\iota \frac{\partial}{\partial x^\kappa} + \check{\alpha}^4_\iota \frac{\partial}{\partial x^4},$$

which together with (12) and (8) gives

$$\frac{\partial}{\partial x'^\iota} + \sigma'_\iota \frac{\partial}{\partial x'_4} = \frac{\partial}{\partial x^\iota} + \sigma_\iota \frac{\partial}{\partial x_4}. \tag{13}$$

A straightforward calculation then shows that the quantity (VIII. 110) which can be written

$$\omega_{\iota\kappa} = \left(\frac{\partial}{\partial x^\iota} + \sigma_\iota \frac{\partial}{\partial x^4}\right)\sigma_\kappa - \left(\frac{\partial}{\partial x^\kappa} + \sigma_\kappa \frac{\partial}{\partial x^4}\right)\sigma_\iota$$

transforms according to the simple law

$$\omega'_{\iota\kappa} = \alpha^4_4 \omega_{\iota\kappa} = \frac{\partial x'^4}{\partial x^4}\omega_{\iota\kappa} \tag{14}$$

under the transformation (5), i.e. each component of the spatial tensor $\omega_{\iota\kappa}$ is multiplied by the ratio between the rates of the coordinate clocks in the two systems.

5. Dual tensors in a three-dimensional space

In a three-dimensional space with a positive definite metric, i.e. $\gamma = |\gamma_{\iota\kappa}| > 0$, the quantities

$$\epsilon_{\iota\kappa\lambda} = \sqrt{\gamma}\,\delta_{\iota\kappa\lambda} \tag{1}$$

are the covariant components of an antisymmetric pseudo-tensor if $\delta_{\iota\kappa\lambda}$ denotes the three-dimensional Levi–Civita symbol defined at the end of § 43. Remembering that the determinant $|\gamma^{\iota\kappa}|$ formed by means of the contravariant components of the metric tensor is equal to γ^{-1}, we find that the contravariant components of the pseudo-tensor (1) are

$$\epsilon^{\iota\kappa\lambda} = \epsilon_{\mu\nu\rho}\gamma^{\mu\iota}\gamma^{\nu\kappa}\gamma^{\rho\lambda} = \sqrt{\gamma}\,\delta_{\mu\nu\rho}\gamma^{\mu\iota}\gamma^{\nu\kappa}\gamma^{\rho\lambda},$$

i.e.

$$\epsilon^{\iota\kappa\lambda} = \sqrt{\gamma}\begin{vmatrix} \gamma^{1\iota} & \gamma^{1\kappa} & \gamma^{1\lambda} \\ \gamma^{2\iota} & \gamma^{2\kappa} & \gamma^{2\lambda} \\ \gamma^{3\iota} & \gamma^{3\kappa} & \gamma^{3\lambda} \end{vmatrix} = \frac{1}{\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}. \tag{2}$$

The covariant components of the axial vector **H** dual to an antisymmetric tensor $H^{\iota\kappa} = -H^{\kappa\iota}$ are now

$$H_\iota = \tfrac{1}{2}\epsilon_{\iota\kappa\lambda}H^{\kappa\lambda} = \tfrac{1}{2}\sqrt{\gamma}\,\delta_{\iota\kappa\lambda}H^{\kappa\lambda}, \tag{3}$$

and the corresponding contravariant components are

$$H^\iota = \tfrac{1}{2}\epsilon^{\iota\kappa\lambda}H_{\kappa\lambda} = \frac{1}{2\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}H_{\kappa\lambda}. \tag{4}$$

For H_ι and H^ι we thus get the following explicit expressions

$$\begin{aligned} (H_1, H_2, H_3) &= \sqrt{\gamma}(H^{23}, H^{31}, H^{12}), \\ (H^1, H^2, H^3) &= \frac{1}{\sqrt{\gamma}}(H_{23}, H_{31}, H_{12}), \end{aligned} \tag{5}$$

which show that the relations reciprocal to (3) and (4) are

$$H^{\iota\kappa} = \epsilon^{\iota\kappa\lambda}H_\lambda, \qquad H_{\iota\kappa} = \epsilon_{\iota\kappa\lambda}H^\lambda. \tag{5'}$$

From two vectors a^ι, b^ι we can build the antisymmetrical tensor $c^{\iota\kappa} = a^\iota b^\kappa - a^\kappa b^\iota$ and the corresponding dual axial vector is the vector product

$$\mathbf{c} = \mathbf{a}\times\mathbf{b} \tag{6}$$

of the two vectors whose components are obtained by means of (5). Similarly to the tensor $\mathrm{curl}_{\iota\kappa}\,\mathbf{a} = \dfrac{\partial a_\kappa}{\partial x^\iota} - \dfrac{\partial a_\iota}{\partial x^\kappa}$ corresponds the axial vector curl **a**.

The tensor dual to an antisymmetrical tensor $V_{\iota\kappa\lambda}$ of rank 3 is a pseudo-scalar

$$V = \frac{1}{3!}\epsilon^{\iota\kappa\lambda}V_{\iota\kappa\lambda} = \frac{1}{\sqrt{\gamma}}\frac{1}{3!}\,\delta_{\iota\kappa\lambda}V_{\iota\kappa\lambda}. \tag{7}$$

If $H^\iota = \dfrac{1}{2\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}H_{\kappa\lambda}$ is the vector dual to the tensor $H_{\iota\kappa}$, the divergence of **H** is dual to the curl of the tensor $H_{\iota\kappa}$, for, according to (X. 75), we have

$$\begin{aligned} \mathrm{div}\,\mathbf{H} &= \frac{1}{\sqrt{\gamma}}\frac{\partial\sqrt{\gamma}H^\iota}{\partial x^\iota} = \frac{1}{2\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}\frac{\partial H_{\kappa\lambda}}{\partial x^\iota} \\ &= \frac{1}{3!}\frac{1}{\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}\left(\frac{\partial H_{\kappa\lambda}}{\partial x^\iota}+\frac{\partial H_{\lambda\iota}}{\partial x^\kappa}+\frac{\partial H_{\iota\kappa}}{\partial x^\lambda}\right) = \frac{1}{3!}\epsilon^{\iota\kappa\lambda}\,\mathrm{curl}_{\iota\kappa\lambda}\{H_{\iota\kappa}\}. \end{aligned} \tag{8}$$

Similarly by means of (5′) and (2) we get

$$\mathrm{div}^\iota\{H^{\iota\kappa}\} = \frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial x^\kappa}(\sqrt{\gamma}H^{\iota\kappa}) = \frac{1}{\sqrt{\gamma}}\,\delta_{\iota\kappa\lambda}\frac{\partial H_\lambda}{\partial x^\kappa} = \frac{1}{2}\epsilon^{\iota\kappa\lambda}\left(\frac{\partial H_\lambda}{\partial x^\kappa}-\frac{\partial H_\kappa}{\partial x^\lambda}\right) = \mathrm{curl}^\iota\mathbf{H}. \tag{9}$$

6. The condition for flat space

A flat space is defined as a space in which it is possible to introduce a system of coordinates which is geodesic at every point. If this condition is fulfilled we obviously have

$$R_{iklm} = 0 \tag{1}$$

at every point, where R_{iklm} is the curvature tensor defined by (IX. 99). Conversely, we shall now show that (1) also represents a sufficient condition for flatness.

A vector field $a^i(x)$ is called *stationary* if it satisfies the equation

$$\frac{\partial a^i}{\partial x^l} = -\Gamma^i_{lr}\, a^r \qquad (2)$$

at every point where $\Gamma^i_{lr}(x)$ is the Christoffel symbol. The equations (2) represent a set of differential equations for the functions $a^i(x)$ which have solutions only if the integrability conditions

$$-\frac{\partial(\Gamma^i_{lr}\, a^r)}{\partial x^m} + \frac{\partial(\Gamma^i_{mr}\, a^r)}{\partial x^l} = \frac{\partial^2 a^i}{\partial x^l \partial x^m} - \frac{\partial^2 a^i}{\partial x^m \partial x^l} = 0 \qquad (3)$$

are satisfied. The conditions (3) may also be written

$$\left(-\frac{\partial \Gamma^i_{kl}}{\partial x^m} + \frac{\partial \Gamma^i_{km}}{\partial x^l} + \Gamma^i_{ls}\,\Gamma^s_{mk} - \Gamma^i_{ms}\,\Gamma^s_{lk}\right) a^k = -R^i{}_{klm}\, a^k = 0. \qquad (3')$$

The integrability conditions are thus fulfilled if (1) holds, and in this case (2) can be integrated and the solution is uniquely determined if the vector $a^i(P)$ at a point P is given. Since

$$da^i = \frac{\partial a^i}{\partial x^l}\, dx^l = -\Gamma^i_{lr}\, a^r\, dx^l,$$

the vectors in a stationary vector field are, according to (IX. 62), obtained by parallel displacements. This is in accordance with (IX. 104), which shows in the case of a vanishing curvature tensor that the result of a parallel displacement of a vector is unique and independent of the path along which the displacement is performed.

Let us now consider four different solutions $a^{(i)}_k = \{a^i_{(1)}, a^i_{(2)}, a^i_{(3)}, a^i_{(4)}\}$ of the differential equations (2) determined by four given vectors $a^i_{(k)}(P)$ at the arbitrary point P. The vectors $a^i_{(k)}(P)$ must be taken to be linearly independent—for instance orthogonal to each other. It can then be shown that we get a transformation $x'^i = x'^i(x^k)$ which leads to a system of coordinates (x'^i) geodesic everywhere if we choose the transformation coefficients

$$\breve{\alpha}^i_k(x) = a^i_{(k)}(x). \qquad (4)$$

In the first place it is seen that the integrability conditions (IX. 14) which also may be written

$$\frac{\partial \breve{\alpha}^i_k}{\partial x'^l} = \frac{\partial \breve{\alpha}^i_l}{\partial x'^k} \qquad (5)$$

are satisfied by the expressions (4). In fact, since the equations (2) hold for each of the vector fields $a^i_{(k)}$, we have

$$\frac{\partial \breve{\alpha}^i_k}{\partial x'^l} = \frac{\partial \breve{\alpha}^i_k}{\partial x^r}\frac{\partial x^r}{\partial x'^l} = \frac{\partial a^i_{(k)}}{\partial x^r}\breve{\alpha}^r_l = -\Gamma^i_{rs}\, a^s_{(k)}\, \breve{\alpha}^r_l = -\Gamma^i_{rs}\, \breve{\alpha}^s_k \breve{\alpha}^r_l, \qquad (6)$$

which on account of the symmetry of Γ^i_{rs} in the lower indices is symmetrical in k and l.

From the Christoffel formulae (IX. 53) and from (6) we now get

$$\Gamma'^i_{kl} = \alpha^i_r\left(\frac{\partial \breve{\alpha}^r_k}{\partial x'^l} + \breve{\alpha}^s_k \breve{\alpha}^t_l \Gamma^r_{st}\right) = \alpha^i_r(-\Gamma^r_{st}\breve{\alpha}^s_k\breve{\alpha}^t_l + \breve{\alpha}^s_k\breve{\alpha}^t_l\Gamma^r_{st}) = 0.$$

Thus the Christoffel symbols—and therefore also $\partial g'_{ik}/\partial x'^l$—vanish at every point in the system of coordinates (x'^i). In this system the components of the metric tensor are thus constants, i.e. we are dealing with a flat space.

7. The action principle and the Hamiltonian equations for a particle in an arbitrary gravitational field

Consider a freely falling particle in an arbitrary external gravitational field For simplicity, we shall make use of the possibility, mentioned in § 113, of introducing a time-orthogonal system of coordinates $(x^i) = (x^\iota, ct)$ in which $\gamma_\iota = g_{\iota 4} = 0$. The connexion between the proper time τ of the particle and the coordinate time t is then given by (VIII. 99), i e.

$$d\tau = dt\,(1+2\chi/c^2-u^2/c^2)^{\frac{1}{2}}, \tag{1}$$

where $$u^2 = \gamma_{\iota\kappa}\,u^\iota u^\kappa = \gamma_{\iota\kappa}\,\dot{x}^\iota \dot{x}^\kappa$$

is the square of the velocity of the particle, and $\chi = \chi(x^\iota, t)$ is the scalar gravitational potential According to (VIII, 85) the time-track of the particle is determined by the variational principle

$$\delta\int\limits_{\tau_1}^{\tau_2} d\tau = \delta\frac{1}{c}\int\sqrt{\left(-g_{ik}\frac{dx^i}{d\lambda}\frac{dx^k}{d\lambda}\right)}\,d\lambda = 0 \tag{2}$$

and, since the integrand is a homogeneous function of the four variables $dx^i/d\lambda$ of first degree, (2) is equivalent to a variational problem with only three dependent variables x^ι and with t as the independent variable.†

Hence the motion of the particle is determined by a variational principle analogous to Hamilton's principle in Newtonian mechanics In fact, multiplying by a constant factor $-\overset{0}{m}_0 c^2$, we get for all variations δx^ι, which vanish for $t = t_1$ and $t = t_2$,

$$\delta\int\limits_{t_1}^{t_2} L(\dot{x}^\iota, x^\iota, t)\,dt = 0, \tag{3}$$

where the Lagrangian L is given by

$$L(\dot{x}^\iota, x^\iota, t) = -\overset{0}{m}_0 c^2\frac{d\tau}{dt} = -\overset{0}{m}_0 c^2(1+2\chi/c^2-\gamma_{\iota\kappa}\dot{x}^\iota\dot{x}^\kappa/c^2)^{\frac{1}{2}} \tag{4}$$

The corresponding Euler equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^\iota}\right)-\frac{\partial L}{\partial x^\iota} = 0 \quad (\iota = 1, 2, 3) \tag{5}$$

are then analogous to the Lagrangian equations of motion in Newtonian mechanics for a system with the generalized coordinates $x^\iota = x^\iota(t)$ From (4) we get

$$\frac{\partial L}{\partial \dot{x}^\iota} = \frac{\overset{0}{m}_0\gamma_{\iota\kappa}\dot{x}^\kappa}{(1+2\chi/c^2-u^2/c^2)^{\frac{1}{2}}} = m u_\iota = p_\iota \tag{6}$$

with m and p_ι given by (X. 12) Further,

$$\frac{\partial L}{\partial x^\iota} = \frac{m}{2}\frac{\partial\gamma_{\kappa\lambda}}{\partial x^\iota}u^\kappa u^\lambda - m\frac{\partial\chi}{\partial x^\iota}.$$

Thus, the Lagrangian equations (5) take the form

$$\frac{d_c p_\iota}{dt} = -m\frac{\partial\chi}{\partial x^\iota} \tag{7}$$

with $d_c p_\iota/dt$ given by (X. 13), which shows that the gravitational force is given

† See, for instance, R Courant and D Hilbert, *Methoden der Mathematischen Physik I*, Berlin 1924, p. 174

by (X. 17) Using the general formula (VIII. 98) for $d\tau/dt$ instead of (VIII. 99) in (4), one could in the same way find the expression for the gravitational force in a general system of coordinates, where $\gamma_i \neq 0$ and, in particular, the formula (X. 18).

From the Lagrangian form (5) of the equations of motion we can now pass over to the canonical Hamiltonian form by the usual procedure. According to (6), the canonically conjugate momenta of the coordinates x^i are the covariant components p_i of the momentum of the particle Thus, by means of (4) and (6), we get for the Hamiltonian function

$$H = p_i \dot{x}^i - L = \mathring{m}_0 c^2(1+2\chi/c^2)(1+2\chi/c^2-u^2/c^2)^{-\frac{1}{2}}$$
$$= mc^2(1+2\chi/c^2), \tag{8}$$

which is identical with the expression (X. 22) for the energy of a particle in a gravitational field with $\gamma_i = 0$. In a static gravitational field, H is a constant of the motion. Solving (6) with respect to u_i, we get

$$u_i = \frac{(c^2+2\chi)^{\frac{1}{2}}}{(\mathring{m}_0^2 c^2+p^2)^{\frac{1}{2}}} p_i, \qquad u^2 = \frac{c^2+2\chi}{\mathring{m}_0^2 c^2+p^2} p^2, \tag{9}$$

which allows us to express H as a function of the canonical variables x^i and p_i. We get

$$H(x^i, p^i) = c(\mathring{m}_0^2 c^2+p^2)^{\frac{1}{2}}(1+2\chi/c^2)^{\frac{1}{2}}. \tag{10}$$

It is now easily verified that the equations (6) and (7) are equivalent to the canonical Hamiltonian equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i} \tag{11}$$

and the whole Hamilton Jacobi theory of integration can then also be applied in this case

In the variation principle (3), the actual motion $x^i = x^i(t)$ is compared with a virtual motion in which the particle has the coordinates $\bar{x}^i = x^i(t)+\delta x^i(t)$ at the time t, i e the symbol δ refers to a variation where the time is kept constant. We shall now consider a more general variation by which the time is varied also, so that the particle in the varied motion has the coordinates $\bar{\bar{x}}^i = x^i(t)+\Delta x^i$ at the time $t+\Delta t$, where Δt may be regarded as an arbitrary infinitesimal function of t. Neglecting terms of order higher than the first, we obviously have

$$\Delta x^i = \delta x^i + \dot{x}^i \Delta t \tag{12}$$

and, consequently,

$$\Delta \dot{x}^i = \Delta \frac{dx^i}{dt} = \frac{\Delta dx^i}{dt} - \frac{dx^i \, \Delta \, dt}{dt^2} = \frac{d\Delta x^i}{dt} - \dot{x}^i \frac{d\Delta t}{dt}$$
$$= \frac{d\delta x^i}{dt} + \ddot{x}^i \Delta t + \dot{x}^i \frac{d\Delta t}{dt} - \dot{x}^i \frac{d\Delta t}{dt},$$

or

$$\Delta \dot{x}^i = \delta \dot{x}^i + \ddot{x}^i \Delta t. \tag{13}$$

Thus, we get for any function $L(x^i, \dot{x}^i, t)$ of the variables x^i, $\dot{x}^i$, and t

$$\Delta L = \frac{\partial L}{\partial x^i}(\delta x^i + \dot{x}^i \Delta t) + \frac{\partial L}{\partial \dot{x}^i}(\delta \dot{x}^i + \ddot{x}^i \Delta t) + \frac{\partial L}{\partial t} \Delta t = \delta L + \frac{dL}{dt} \Delta t. \tag{14}$$

Hence,

$$\Delta(L\, dt) = \Delta L\, dt + L\, d\Delta t = \delta L\, dt + \frac{d(L\, \Delta t)}{dt} dt \tag{15}$$

and

$$\Delta \int_{t_1}^{t_2} L\, dt = \delta \int_{t_1}^{t_2} L\, dt + (L\, \Delta t) \Big|_{t_1}^{t_2}. \tag{16}$$

According to (3) and (16), we thus have

$$\Delta \int_{t_1}^{t_2} L\,dt = 0 \tag{17}$$

provided that we choose

$$\Delta t = 0 \quad \text{for } t = t_1 \text{ and } t = t_2. \tag{18}$$

On account of (8) this may also be written

$$\Delta \int p_i\,dx^i - \Delta \int H\,dt = 0. \tag{19}$$

Now we have

$$\Delta(H\,dt) = H\Delta\,dt + \Delta H\,dt = H\,d\Delta t + \left(\delta H + \frac{dH}{dt}\Delta t\right)dt$$

$$= \frac{d(H\Delta t)}{dt}\,dt + \delta H\,dt, \tag{20}$$

where we have made use of (14) applied to the function H.

Since the Δ-variation is more general than the δ-variation, we may now impose the further condition

$$\delta H = 0, \tag{21}$$

which means that the energy of the particle at any time has the same value in the actual and the virtual motion. This does not necessarily mean that the energy must be constant along the path. From (20) and (21) we then get

$$\Delta \int_{t_1}^{t_2} H\,dt = \int_{t_1}^{t_2} \Delta(H\,dt) = \int_{t_1}^{t_2} \frac{d(H\Delta t)}{dt}\,dt = (H\Delta t)\Big|_{t_1}^{t_2} = 0$$

and (19) reduces to

$$\Delta \int_{t_1}^{t_2} p_i\,dx^i = 0 \tag{22}$$

(22) together with (21) and (18) correspond to Maupertuis's principle of least action in its most general form

For a static field, the energy H is equal to a constant E along the actual path, and Maupertuis's principle states that the 'action' $\int_{t_1}^{t_2} p_i\,dx^i$ in first approximation has the same value for all virtual motions with the same energy E. In the static case, the time may even be eliminated entirely from (22). Since $p_i = mu_i$ and $d\sigma/dt = u$, we have

$$p_i\,dx^i = p_i u^i\,dt = mu^2\,d\sigma/u = mu\,d\sigma = p\,d\sigma$$

and (22) may be written

$$\Delta \int_{(x_1^i)}^{(x_2^i)} p\,d\sigma = 0, \tag{23}$$

where (x_1^i) and (x_2^i) are the coordinates of the particle at the times t_1 and t_2, respectively.

Putting $H = E$ in (10), and solving with respect to p, we get

$$p = \{(E/c)^2 - m_0^2c^2 - 2m_0^2\chi\}^{\frac{1}{2}}(1+2\chi/c^2)^{-\frac{1}{2}}. \tag{24}$$

Hence, the variational principle takes the form

$$\Delta \int_{(x_1^i)}^{(x_2^i)} (1+2\chi/c^2)^{-\frac{1}{2}}\{(E/c)^2 - m_0^2c^2 - 2m_0^2\chi\}^{\frac{1}{2}}\,d\sigma = 0, \tag{25}$$

where E as well as the end points (x_1^ι) and (x_2^ι) are to be kept constant by the Δ-variation. In this form, the variation principle is analogous to Hertz's 'principle of the straightest path' in Newtonian mechanics. In the limit of weak fields and small velocities the integrand in (25) reduces to

$$\sqrt{\{2\mathring{m}_0(\epsilon-\mathring{m}_0\chi)\}}, \quad \text{where } \epsilon = E-\mathring{m}_0 c^2,$$

which is the integrand occurring in Hertz's principle.

If $\chi = 0$, the integrand in (25) is a constant and we get

$$\delta \int d\sigma = 0, \tag{26}$$

which shows that the path is a geodesic in the space defined by the metric tensor $\gamma_{\iota\kappa}$.

As remarked by Hertz,† the more general case, where $\chi(x^\iota) \neq 0$, can formally be treated also as a problem of determining a straightest path, but in a space where the line element is defined by

$$d\Sigma = \Gamma_{\iota\kappa}\, dx^\iota\, dx^\kappa = p^2\, d\sigma^2,$$

i.e.

$$\Gamma_{\iota\kappa} = (1+2\chi/c^2)^{-1}\{(E/c)^2-\mathring{m}_0^2 c^2-2\mathring{m}_0^2\chi\}\gamma_{\iota\kappa}. \tag{27}$$

Such a treatment is, however, quite formal, since the geometrical structure of the space as determined by natural measuring sticks is described by $\gamma_{\iota\kappa}$ and not by $\Gamma_{\iota\kappa}$.

8. The connexion between the determinants of the space-time metric tensor and the spatial metric tensor

According to (VIII. 64, 63) we have

$$\gamma_{\iota\kappa} = g_{\iota\kappa}+\gamma_\iota\gamma_\kappa = g_{\iota\kappa}-\frac{g_{\iota 4}g_{\kappa 4}}{g_{44}}. \tag{1}$$

The determinant γ may be written in the form

$$\gamma = \begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{vmatrix} = \frac{1}{g_{44}} \begin{vmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & g_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & g_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & g_{34} \\ 0 & 0 & 0 & g_{44} \end{vmatrix}.$$

Since a determinant is unchanged if we multiply the elements of one of the columns by a common factor and add them to the corresponding elements of another column we have

$$g_{44}\gamma = \begin{vmatrix} \gamma_{11}+\dfrac{g_{14}g_{14}}{g_{44}} & \gamma_{12} & \gamma_{13} & g_{14} \\ \gamma_{21}+\dfrac{g_{24}g_{14}}{g_{44}} & \gamma_{22} & \gamma_{23} & g_{24} \\ \gamma_{31}+\dfrac{g_{34}g_{14}}{g_{44}} & \gamma_{32} & \gamma_{33} & g_{34} \\ \dfrac{g_{44}g_{14}}{g_{44}} & 0 & 0 & g_{44} \end{vmatrix} = \begin{vmatrix} g_{11} & \gamma_{12} & \gamma_{13} & g_{14} \\ g_{21} & \gamma_{22} & \gamma_{23} & g_{24} \\ g_{31} & \gamma_{32} & \gamma_{33} & g_{34} \\ g_{41} & 0 & 0 & g_{44} \end{vmatrix}$$

on account of (1).

† H. Hertz, *Die Prinzipien der Mechanik*, 2nd ed., Leipzig (1910).

If we apply the same procedure to the second and the third column we get

$$g_{44}\gamma = \begin{vmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{vmatrix} = g \tag{2}$$

Since both g_{44} and g are negative, γ must thus be positive and we have

$$\sqrt{(-g)} = \sqrt{(-g_{44})}\,\sqrt{\gamma} = \sqrt{\gamma}\,\sqrt{(1+2\chi/c^2)}. \tag{3}$$

9. The derivatives of the function $\mathfrak{L}$ with respect to g^{lm}_k and g^{lm} and some identities containing these derivatives

According to (XI 132) the function $\mathfrak{L}$ can be written

$$\mathfrak{L} = \sqrt{(-g)}(A-B), \tag{1}$$

where
$$A = g^{ik}\Gamma^t_{ik}\Gamma^s_{ts}, \qquad B = g^{lr}\Gamma^k_{lm}\Gamma^m_{rk} \tag{2}$$

Hence, since g^{ik} and g_{ik} are kept constant in partial derivation with respect to g^{lm}_k we have

$$\frac{\partial\mathfrak{L}}{\partial g^{lm}_k} = \sqrt{(-g)}\left(\frac{\partial A}{\partial g^{lm}_k} - \frac{\partial B}{\partial g^{lm}_k}\right), \tag{3}$$

$$\frac{\partial A}{\partial g^{lm}_k} = g^{ir}\Gamma^t_{ir}\frac{\partial\Gamma^s_{ts}}{\partial g^{lm}_k} + \Gamma^s_{rs}\frac{\partial(g^{it}\Gamma^r_{it})}{\partial g^{lm}_k}. \tag{4}$$

From (IX 68′) we get
$$\Gamma^s_{ts} = -\tfrac{1}{2}g_{lm}\,g^{lm}_t, \tag{5}$$

i.e.
$$\frac{\partial\Gamma^s_{ts}}{\partial g^{lm}_k} = -\tfrac{1}{2}g_{lm}\,\delta^k_t \tag{6}$$

Further, by contraction of (IX. 68) with respect to the indices k and l

$$g^{it}\Gamma^r_{it} = -\frac{\partial g^{rm}}{\partial x^m} - g^{rk}\Gamma^t_{kt} = -\tfrac{1}{2}(g^{rm}_m + g^{mr}_m) + \tfrac{1}{2}g^{rk}g_{lm}\,g^{lm}_k. \tag{7}$$

Hence
$$\frac{\partial(g^{it}\Gamma^r_{it})}{\partial g^{lm}_k} = -\tfrac{1}{2}(\delta^r_l\,\delta^k_m + \delta^r_m\,\delta^k_l) + \tfrac{1}{2}g^{rk}g_{lm} \tag{8}$$

which is symmetrical in l and m (cf. the remarks following (XI 140)).

Thus, by means of (4)–(8), we get

$$\frac{\partial A}{\partial g^{lm}_k} = -\tfrac{1}{2}g^{rs}\Gamma^k_{rs}g_{lm} + \Gamma^s_{rs}[-\tfrac{1}{2}(\delta^r_l\,\delta^k_m + \delta^r_m\,\delta^k_l) + \tfrac{1}{2}g^{rk}g_{lm}] \tag{9}$$

In order to find the corresponding derivative of B we consider a variation of the variables g^{lm}_k for constant g^{ik}. The corresponding variation of B is

$$\begin{aligned}\delta B &= g^{lr}(\Gamma^k_{lm}\,\delta\Gamma^m_{rk} + \Gamma^m_{rk}\delta\Gamma^k_{lm}) \\ &= \Gamma^k_{lm}\delta(g^{lr}\Gamma^m_{rk}) + \Gamma^m_{rk}\delta(g^{rl}\Gamma^k_{lm}).\end{aligned}$$

If, in the last term, we perform a cyclic permutation of the four summation indices k, l, r, m we get

$$\delta B = \Gamma^k_{lm}\delta(g^{lr}\Gamma^m_{rk} + g^{mr}\Gamma^l_{rk}),$$

which by means of (IX 68) reduces to

$$\delta B = -\Gamma^k_{lm}\delta g^{lm}_k.$$

Hence
$$\frac{\partial B}{\partial g^{lm}_k} = -\Gamma^k_{lm}, \tag{10}$$

which together with (9) and (3) leads to the desired formula

$$\frac{\partial \mathfrak{L}}{\partial g_k^{lm}} = \sqrt{(-g)}\{\Gamma^k_{lm} - \tfrac{1}{2}(\delta^k_l \Gamma^r_{mr} + \delta^k_m \Gamma^r_{lr}) - \tfrac{1}{2}(g^{rs}\Gamma^k_{rs} - g^{rk}\Gamma^s_{rs})g_{lm}\}. \tag{11}$$

In the same way we can find the derivative of $\mathfrak{L}$ with respect to g^{lm}. However, it is easier to use the equation (XI 139) according to which

$$\frac{\partial \mathfrak{L}}{\partial g^{lm}} = \sqrt{(-g)}(R_{lm} - \tfrac{1}{2}Rg_{lm}) + \frac{\partial}{\partial x^k}\left(\frac{\partial \mathfrak{L}}{\partial g_k^{lm}}\right). \tag{12}$$

The last term may be obtained from (11) by differentiation with respect to x^k while the first term is given by (IX. 114, 111) Besides the components of the metric tensor and Christoffel symbols these expressions contain derivatives of the Christoffel symbols. Since $\mathfrak{L}$ and therefore also the left-hand side of (12) does not contain derivatives of Christoffel symbols, all terms on the right-hand side containing such derivatives must cancel and we may therefore omit them from the beginning in our calculation. By differentiation of (11) with respect to x^k we are then left with terms containing derivatives of $\sqrt{(-g)}$, g^{rs}, g_{lm}, only By means of (IX 69′, 68, 51) these terms and therefore also the right-hand side of (12) may be expressed in terms of components of the metric tensor and of Christoffel symbols After a straightforward calculation one finds

$$\frac{\partial \mathfrak{L}}{\partial g^{lm}} = \sqrt{(-g)}\{\Gamma^r_{ls}\Gamma^s_{mr} - \Gamma^r_{lr}\Gamma^s_{ms} - \tfrac{1}{2}(g^{rs}\Gamma^k_{rs} - g^{rk}\Gamma^s_{rs})(\Gamma_{m,lk} + \Gamma_{l,mk})\} - \tfrac{1}{2}g_{lm}\mathfrak{L}. \tag{13}$$

The expressions (11) and (13) for the derivatives of $\mathfrak{L}$ with respect to g_k^{lm} and g^{lm} are easily seen to be in accordance with the equations (XI 141, 142).

If we now substitute (11) and (13) in the last term $\mathfrak{L}_i^k$ on the right-hand side of (XI. 161) and use the equation

$$g_m^{kl} = -(\Gamma^k_{mr}g^{rl} + \Gamma^l_{mr}g^{kr})$$

following from (IX 68), we get for $\mathfrak{L}_i^k$ an expression containing a large number of terms, which, however, cancel in pairs. The final result is therefore that the quantities $\mathfrak{L}_i^k$ are identically zero

$$\mathfrak{L}_i^k = \frac{1}{\kappa}\left(\frac{1}{2}\frac{\partial \mathfrak{L}}{\partial g_k^{lm}}g_i^{lm} - \frac{\partial \mathfrak{L}}{\partial g_m^{il}}g_m^{kl} - \frac{\partial \mathfrak{L}}{\partial g^{il}}g^{kl} - \tfrac{1}{2}\delta_i^k \mathfrak{L}\right) = 0. \tag{14}$$

AUTHOR INDEX

SUBJECT INDEX