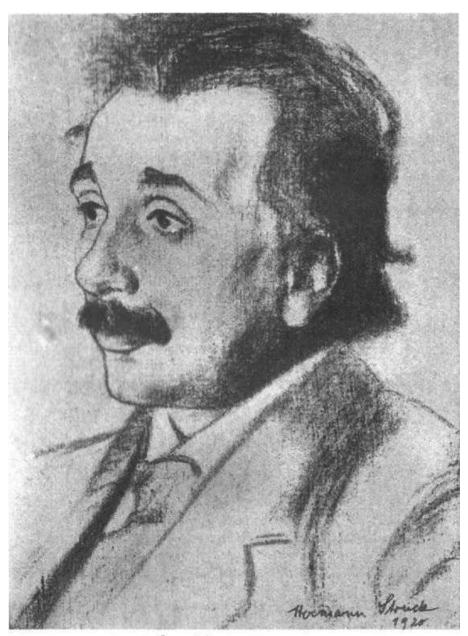
RELATIVITY: THE SPECIAL AND GENERAL THEORY

Albert Einstein

RELATIVITY

THE SPECIAL AND GENERAL THEORY

ALBERT EINSTEIN



A. Vintin

RELATIVITY

THE SPECIAL AND GENERAL THEORY

 \mathbf{BY}

ALBERT EINSTEIN, Ph.D.

PROFESSOR OF PHYSICS IN THE UNIVERSITY OF BERLIN

TRANSLATED BY

ROBERT W. LAWSON, D.Sc., F. Inst. P.

UNIVERSITY OF SHEFFIELD

1920 EDITION

COPYRIGHT INFORMATION

Book: Relativity: The Special and General Theory

Author: Albert Einstein, 1879–1955

First published: 1920

This PDF file contains the text of the first English translation of $\ddot{U}ber\ die\ spezielle\ und\ die\ allgemeine\ Relativitätstheorie$, published in 1920. (The index has not been included. A few misprints in the original text have been corrected. They are marked by footnotes enclosed in square brackets and signed "J.M.") The original book is in the public domain in the United States. However, since Einstein died in 1955, it is still under copyright in most other countries, for example, those that use the life of the author + 50 years or life + 70 years for the duration of copyright. Readers outside the United States should check their own countries' copyright laws to be certain they can legally download this ebook.

PREFACE

HE present book is intended, as far as possible, to give an exact insight into the theory of Relativity to those readers who, from a general scientific and philosophical point of view, are interested in the theory, but who are not conversant with the mathematical apparatus ¹ of theoretical physics. The work presumes a standard of education corresponding to that of a university matriculation examination, and, despite the shortness of the book, a fair amount of patience and force of will on the part of the reader. The author has spared himself no pains in his endeavour to present the main ideas in the simplest and most intelligible form, and on the whole, in the sequence and connection in which they actually originated. In the interest of clearness, it appeared to me inevitable that I should repeat myself frequently, without paying the slightest attention to the elegance of the presentation. I adhered scrupulously to the precept of that brilliant theoretical physicist, L. Boltzmann, according to whom matters of elegance ought to be left to the tailor and to the cobbler. I make no pretence of having withheld from the reader difficulties which are inherent to the subject. On the other hand, I have purposely treated the empirical physical foundations of the theory in a "step-motherly" fashion, so that readers unfamiliar with physics may not feel like the wanderer who was unable to see

¹ The mathematical fundaments of the special theory of relativity are to be found in the original papers of H. A. Lorentz, A. Einstein, H. Minkowski published under the title *Das Relativitätsprinzip* (The Principle of Relativity) in B. G. Teubner's collection of monographs *Fortschritte der mathematischen Wissenschaften* (Advances in the Mathematical Sciences), also in M. Laue's exhaustive book *Das Relativitäts prinzip*—published by Friedr. Vieweg & Son, Braunschweig. The general theory of relativity, together with the necessary parts of the theory of invariants, is dealt with in the author's book *Die Grundlagen der allgemeinen Relativitätstheorie* (The Foundations of the General Theory of Relativity)—Joh. Ambr. Barth, 1916; this book assumes some familiarity with the special theory of relativity.

the forest for trees. May the book bring some one a few happy hours of suggestive thought!

A. EINSTEIN

December, 1916

NOTE TO THE THIRD EDITION

N the present year (1918) an excellent and detailed manual on the general theory of relativity, written by H. Weyl, was published by the firm Julius Springer (Berlin). This book, entitled *Raum—Zeit—Materie* (Space—Time—Matter), may be warmly recommended to mathematicians and physicists.

BIOGRAPHICAL NOTE

LBERT EINSTEIN is the son of German-Jewish parents. He was born in 1879 in the town of Ulm, Würtemberg, Germany. His schooldays were spent in Munich, where he attended the *Gymnasium* until his sixteenth year. After leaving school at Munich, he accompanied his parents to Milan, whence he proceeded to Switzerland six months later to continue his studies.

From 1896 to 1900 Albert Einstein studied mathematics and physics at the Technical High School in Zurich, as he intended becoming a secondary school (*Gymnasium*) teacher. For some time afterwards he was a private tutor, and having meanwhile become naturalised, he obtained a post as engineer in the Swiss Patent Office in 1902, which position he occupied till 1909. The main ideas involved in the most important of Einstein's theories date back to this period. Amongst these may be mentioned: *The Special Theory of Relativity, Inertia of Energy, Theory of the Brownian Movement*, and the *Quantum-Law of the Emission and Absorption of Light* (1905). These were followed some years later by the *Theory of the Specific Heat of Solid Bodies*, and the fundamental idea of the *General Theory of Relativity*.

During the interval 1909 to 1911 he occupied the post of Professor *Extraordinarius* at the University of Zurich, afterwards being appointed to the University of Prague, Bohemia, where he remained as Professor *Ordinarius* until 1912. In the latter year Professor Einstein accepted a similar chair at the *Polytechnikum*, Zurich, and continued his activities there until 1914, when he received a call to the Prussian Academy of Science, Berlin, as successor to Van't Hoff. Professor Einstein is able to devote himself freely to his studies at the Berlin Academy, and it was here that he

succeeded in completing his work on the *General Theory of Relativity* (1915–17). Professor Einstein also lectures on various special branches of physics at the University of Berlin, and, in addition, he is Director of the Institute* for Physical Research of the *Kaiser Wilhelm Gesellschaft*.

Professor Einstein has been twice married. His first wife, whom he married at Berne in 1903, was a fellow-student from Serbia. There were two sons of this marriage, both of whom are living in Zurich, the elder being sixteen years of age. Recently Professor Einstein married a widowed cousin, with whom he is now living in Berlin.

R. W. L.

^{[*} This word was misprinted Institute in the original book.—J.M.]

TRANSLATOR'S NOTE

In presenting this translation to the English-reading public, it is hardly necessary for me to enlarge on the Author's prefatory remarks, except to draw attention to those additions to the book which do not appear in the original.

At my request, Professor Einstein kindly supplied me with a portrait of himself, by one of Germany's most celebrated artists. Appendix III, on "The Experimental Confirmation of the General Theory of Relativity," has been written specially for this translation. Apart from these valuable additions to the book, I have included a biographical note on the Author, and, at the end of the book, an Index and a list of English references to the subject. This list, which is more suggestive than exhaustive, is intended as a guide to those readers who wish to pursue the subject farther.

I desire to tender my best thanks to my colleagues Professor S. R. Milner, D.Sc., and Mr. W. E. Curtis, A.R.C.Sc., F.R.A.S., also to my friend Dr. Arthur Holmes, A.R.C.Sc., F.G.S., of the Imperial College, for their kindness in reading through the manuscript, for helpful criticism, and for numerous suggestions. I owe an expression of thanks also to Messrs. Methuen for their ready counsel and advice, and for the care they have bestowed on the work during the course of its publication.

ROBERT W. LAWSON

THE PHYSICS LABORATORY
THE UNIVERSITY OF SHEFFIELD
June 12, 1920

CONTENTS

PART I

THE SPECIAL THEORY OF RELATIVITY

- I. Physical Meaning of Geometrical Propositions
- II. The System of Co-ordinates
- III. Space and Time in Classical Mechanics
- IV. The Galileian System of Co-ordinates
- V. The Principle of Relativity (in the Restricted Sense)
- VI. The Theorem of the Addition of Velocities employed in Classical Mechanics
- VII. The Apparent Incompatibility of the Law of Propagation of Light with the Principle of Relativity
- VIII. On the Idea of Time in Physics
 - IX. The Relativity of Simultaneity
 - X. On the Relativity of the Conception of Distance
 - XI. The Lorentz Transformation
- XII. The Behaviour of Measuring-Rods and Clocks in Motion
- XIII. Theorem of the Addition of Velocities. The Experiment of Fizeau
- XIV. The Heuristic Value of the Theory of Relativity
- XV. General Results of the Theory
- XVI. Experience and the Special Theory of Relativity
- XVII. Minkowski's Four-dimensional Space

PART II

THE GENERAL THEORY OF RELATIVITY

- XVIII. Special and General Principle of Relativity
 - XIX. The Gravitational Field
 - XX. The Equality of Inertial and Gravitational Mass as an Argument for the General Postulate of Relativity
 - XXI. In what Respects are the Foundations of Classical Mechanics and of the Special Theory of Relativity unsatisfactory?
- XXII. A Few Inferences from the General Principle of Relativity
- XXIII. Behaviour of Clocks and Measuring-Rods on a Rotating Body of Reference
- XXIV. Euclidean and Non-Euclidean Continuum
- XXV. Gaussian Co-ordinates
- XXVI. The Space-time Continuum of the Special Theory of Relativity considered as a Euclidean Continuum
- XXVII. The Space-time Continuum of the General Theory of Relativity is not a Euclidean Continuum
- XXVIII. Exact Formulation of the General Principle of Relativity
 - XXIX. The Solution of the Problem of Gravitation on the Basis of the General Principle of Relativity

PART III

CONSIDERATIONS ON THE UNIVERSE AS A WHOLE

- XXX. Cosmological Difficulties of Newton's Theory
- XXXI. The Possibility of a "Finite" and yet "Unbounded" Universe
- XXXII. The Structure of Space according to the General Theory of Relativity

APPENDICES

- I. Simple Derivation of the Lorentz Transformation
- II. Minkowski's Four-dimensional Space ("World") [Supplementary to Section XVII.]
- III. The Experimental Confirmation of the General Theory of Relativity
 - (a) Motion of the Perihelion of Mercury
 - (b) Deflection of Light by a Gravitational Field
 - (c) Displacement of Spectral Lines towards the Red

BIBLIOGRAPHY

RELATIVITY

PART I THE SPECIAL THEORY OF RELATIVITY

I

PHYSICAL MEANING OF GEOMETRICAL PROPOSITIONS

In your schooldays most of you who read this book made acquaintance with the noble building of Euclid's geometry, and you remember—perhaps with more respect than love—the magnificent structure, on the lofty staircase of which you were chased about for uncounted hours by conscientious teachers. By reason of your past experience, you would certainly regard every one with disdain who should pronounce even the most out-of-the-way proposition of this science to be untrue. But perhaps this feeling of proud certainty would leave you immediately if some one were to ask you: "What, then, do you mean by the assertion that these propositions are true?" Let us proceed to give this question a little consideration.

Geometry sets out from certain conceptions such as "plane," "point," and "straight line," with which we are able to associate more or less definite ideas, and from certain simple propositions (axioms) which, in virtue of these ideas, we are inclined to accept as "true." Then, on the basis of a logical process, the justification of which we feel ourselves compelled to admit, all remaining propositions are shown to follow from those axioms, *i.e.* they are proven. A proposition is then correct ("true") when it has been derived in the recognised manner from the axioms. The question of the "truth" of the individual geometrical propositions is thus reduced to one of the

"truth" of the axioms. Now it has long been known that the last question is not only unanswerable by the methods of geometry, but that it is in itself entirely without meaning. We cannot ask whether it is true that only one straight line goes through two points. We can only say that Euclidean geometry deals with things called "straight lines," to each of which is ascribed the property of being uniquely determined by two points situated on it. The concept "true" does not tally with the assertions of pure geometry, because by the word "true" we are eventually in the habit of designating always the correspondence with a "real" object; geometry, however, is not concerned with the relation of the ideas involved in it to objects of experience, but only with the logical connection of these ideas among themselves.

It is not difficult to understand why, in spite of this, we feel constrained to call the propositions of geometry "true." Geometrical ideas correspond to more or less exact objects in nature, and these last are undoubtedly the exclusive cause of the genesis of those ideas. Geometry ought to refrain from such a course, in order to give to its structure the largest possible logical unity. The practice, for example, of seeing in a "distance" two marked positions on a practically rigid body is something which is lodged deeply in our habit of thought. We are accustomed further to regard three points as being situated on a straight line, if their apparent positions can be made to coincide for observation with one eye, under suitable choice of our place of observation.

If, in pursuance of our habit of thought, we now supplement the propositions of Euclidean geometry by the single proposition that two points on a practically rigid body always correspond to the same distance (line-interval), independently of any changes in position to which we may subject the body, the propositions of Euclidean geometry then resolve themselves into propositions on the possible relative position of practically rigid bodies. Geometry which has

¹ It follows that a natural object is associated also with a straight line. Three points A, B and C on a rigid body thus lie in a straight line when, the points A and C being given, B is chosen such that the sum of the distances AB and BC is as short as possible. This incomplete suggestion will suffice for our present purpose.

been supplemented in this way is then to be treated as a branch of physics. We can now legitimately ask as to the "truth" of geometrical propositions interpreted in this way, since we are justified in asking whether these propositions are satisfied for those real things we have associated with the geometrical ideas. In less exact terms we can express this by saying that by the "truth" of a geometrical proposition in this sense we understand its validity for a construction with ruler and compasses.

Of course the conviction of the "truth" of geometrical propositions in this sense is founded exclusively on rather incomplete experience. For the present we shall assume the "truth" of the geometrical propositions, then at a later stage (in the general theory of relativity) we shall see that this "truth" is limited, and we shall consider the extent of its limitation.

II

THE SYSTEM OF CO-ORDINATES

N the basis of the physical interpretation of distance which has been indicated, we are also in a position to establish the distance between two points on a rigid body by means of measurements. For this purpose we require a "distance" (rod S) which is to be used once and for all, and which we employ as a standard measure. If, now, A and B are two points on a rigid body, we can construct the line joining them according to the rules of geometry; then, starting from A, we can mark off the distance S time after time until we reach B. The number of these operations required is the numerical measure of the distance AB. This is the basis of all measurement of length.¹

Every description of the scene of an event or of the position of an object in space is based on the specification of the point on a rigid body (body of reference) with which that event or object coincides. This applies not only to scientific description, but also to everyday life. If I analyse the place specification "Trafalgar Square, London," ² I arrive at the following result. The earth is the rigid body to which the specification of place refers; "Trafalgar Square, London" is a

¹ Here we have assumed that there is nothing left over, *i.e.* that the measurement gives a whole number. This difficulty is got over by the use of divided measuring-rods, the introduction of which does not demand any fundamentally new method.

² I have chosen this as being more familiar to the English reader than the "Potsdamer Platz, Berlin," which is referred to in the original. (R. W. L.)

well-defined point, to which a name has been assigned, and with which the event coincides in space.¹

This primitive method of place specification deals only with places on the surface of rigid bodies, and is dependent on the existence of points on this surface which are distinguishable from each other. But we can free ourselves from both of these limitations without altering the nature of our specification of position. If, for instance, a cloud is hovering over Trafalgar Square, then we can determine its position relative to the surface of the earth by erecting a pole perpendicularly on the Square, so that it reaches the cloud. The length of the pole measured with the standard measuring-rod, combined with the specification of the position of the foot of the pole, supplies us with a complete place specification. On the basis of this illustration, we are able to see the manner in which a refinement of the conception of position has been developed.

- (a) We imagine the rigid body, to which the place specification is referred, supplemented in such a manner that the object whose position we require is reached by the completed rigid body.
- (b) In locating the position of the object, we make use of a number (here the length of the pole measured with the measuring-rod) instead of designated points of reference.
- (c) We speak of the height of the cloud even when the pole which reaches the cloud has not been erected. By means of optical observations of the cloud from different positions on the ground, and taking into account the properties of the propagation of light, we determine the length of the pole we should have required in order to reach the cloud.

From this consideration we see that it will be advantageous if, in the description of position, it should be possible by means of numerical measures to make ourselves independent of the existence of marked positions (possessing names) on the rigid body of reference. In the physics of measurement this is attained by the application of the Cartesian system of co-ordinates.

¹ It is not necessary here to investigate further the significance of the expression "coincidence in space." This conception is sufficiently obvious to ensure that differences of opinion are scarcely likely to arise as to its applicability in practice.

This consists of three plane surfaces perpendicular to each other and rigidly attached to a rigid body. Referred to a system of coordinates, the scene of any event will be determined (for the main part) by the specification of the lengths of the three perpendiculars or co-ordinates (x, y, z) which can be dropped from the scene of the event to those three plane surfaces. The lengths of these three perpendiculars can be determined by a series of manipulations with rigid measuring-rods performed according to the rules and methods laid down by Euclidean geometry.

In practice, the rigid surfaces which constitute the system of coordinates are generally not available; furthermore, the magnitudes of the co-ordinates are not actually determined by constructions with rigid rods, but by indirect means. If the results of physics and astronomy are to maintain their clearness, the physical meaning of specifications of position must always be sought in accordance with the above considerations.¹

We thus obtain the following result: Every description of events in space involves the use of a rigid body to which such events have to be referred. The resulting relationship takes for granted that the laws of Euclidean geometry hold for "distances," the "distance" being represented physically by means of the convention of two marks on a rigid body.

¹ A refinement and modification of these views does not become necessary until we come to deal with the general theory of relativity, treated in the second part of this book.

III

SPACE AND TIME IN CLASSICAL MECHANICS

"HE purpose of mechanics is to describe how bodies change their position in space with time." I should load my conscience with grave sins against the sacred spirit of lucidity were I to formulate the aims of mechanics in this way, without serious reflection and detailed explanations. Let us proceed to disclose these sins.

It is not clear what is to be understood here by "position" and "space." I stand at the window of a railway carriage which is travelling uniformly, and drop a stone on the embankment, without throwing it. Then, disregarding the influence of the air resistance, I see the stone descend in a straight line. A pedestrian who observes the misdeed from the footpath notices that the stone falls to earth in a parabolic curve. I now ask: Do the "positions" traversed by the stone lie "in reality" on a straight line or on a parabola? Moreover, what is meant here by motion "in space"? From the considerations of the previous section the answer is self-evident. In the first place, we entirely shun the vague word "space," of which, we must honestly acknowledge, we cannot form the slightest conception, and we replace it by "motion relative to a practically rigid body of reference." The positions relative to the body of reference (railway carriage or embankment) have already been defined in detail in the preceding section. If instead of "body of reference" we insert "system of coordinates," which is a useful idea for mathematical description, we are in a position to say: The stone traverses a straight line relative to a system of co-ordinates rigidly attached to the carriage, but relative to a system of co-ordinates rigidly attached to the ground (embankment)

it describes a parabola. With the aid of this example it is clearly seen that there is no such thing as an independently existing trajectory (lit. "path-curve" ¹), but only a trajectory relative to a particular body of reference.

In order to have a *complete* description of the motion, we must specify how the body alters its position with time; i.e. for every point on the trajectory it must be stated at what time the body is situated there. These data must be supplemented by such a definition of time that, in virtue of this definition, these time-values can be regarded essentially as magnitudes (results of measurements) capable of observation. If we take our stand on the ground of classical mechanics, we can satisfy this requirement for our illustration in the following manner. We imagine two clocks of identical construction; the man at the railway-carriage window is holding one of them, and the man on the footpath the other. Each of the observers determines the position on his own reference-body occupied by the stone at each tick of the clock he is holding in his hand. In this connection we have not taken account of the inaccuracy involved by the finiteness of the velocity of propagation of light. With this and with a second difficulty prevailing here we shall have to deal in detail later.

¹ That is, a curve along which the body moves.

IV

THE GALILEIAN SYSTEM OF CO-ORDINATES

S is well known, the fundamental law of the mechanics of Galilei-Newton, which is known as the law of inertia, can be stated thus: A body removed sufficiently far from other bodies continues in a state of rest or of uniform motion in a straight line. This law not only says something about the motion of the bodies, but it also indicates the reference-bodies or systems of co-ordinates, permissible in mechanics, which can be used in mechanical description. The visible fixed stars are bodies for which the law of inertia certainly holds to a high degree of approximation. Now if we use a system of co-ordinates which is rigidly attached to the earth, then, relative to this system, every fixed star describes a circle of immense radius in the course of an astronomical day, a result which is opposed to the statement of the law of inertia. So that if we adhere to this law we must refer these motions only to systems of co-ordinates relative to which the fixed stars do not move in a circle. A system of co-ordinates of which the state of motion is such that the law of inertia holds relative to it is called a "Galileian system of coordinates." The laws of the mechanics of Galilei-Newton can be regarded as valid only for a Galileian system of co-ordinates.

\mathbf{V}

THE PRINCIPLE OF RELATIVITY (IN THE RESTRICTED SENSE)

N order to attain the greatest possible clearness, let us return to our example of the railway carriage supposed to be travelling uniformly. We call its motion a uniform translation ("uniform" because it is of constant velocity and direction, "translation" because although the carriage changes its position relative to the embankment yet it does not rotate in so doing). Let us imagine a raven flying through the air in such a manner that its motion, as observed from the embankment, is uniform and in a straight line. If we were to observe the flying raven from the moving railway carriage, we should find that the motion of the raven would be one of different velocity and direction, but that it would still be uniform and in a straight line. Expressed in an abstract manner we may say: If a mass m is moving uniformly in a straight line with respect to a co-ordinate system K, then it will also be moving uniformly and in a straight line relative to a second co-ordinate system K', provided that the latter is executing a uniform translatory motion with respect to K. In accordance with the discussion contained in the preceding section, it follows that:

If K is a Galileian co-ordinate system, then every other co-ordinate system K' is a Galileian one, when, in relation to K, it is in a condition of uniform motion of translation. Relative to K' the mechanical laws of Galilei-Newton hold good exactly as they do with respect to K.

We advance a step farther in our generalisation when we express the tenet thus: If, relative to K, K' is a uniformly moving co-ordinate system devoid of rotation, then natural phenomena run their course with respect to K' according to exactly the same general laws as with respect to K. This statement is called the *principle of relativity* (in the restricted sense).

As long as one was convinced that all natural phenomena were capable of representation with the help of classical mechanics, there was no need to doubt the validity of this principle of relativity. But in view of the more recent development of electrodynamics and optics it became more and more evident that classical mechanics affords an insufficient foundation for the physical description of all natural phenomena. At this juncture the question of the validity of the principle of relativity became ripe for discussion, and it did not appear impossible that the answer to this question might be in the negative.

Nevertheless, there are two general facts which at the outset speak very much in favour of the validity of the principle of relativity. Even though classical mechanics does not supply us with a sufficiently broad basis for the theoretical presentation of all physical phenomena, still we must grant it a considerable measure of "truth," since it supplies us with the actual motions of the heavenly bodies with a delicacy of detail little short of wonderful. The principle of relativity must therefore apply with great accuracy in the domain of *mechanics*. But that a principle of such broad generality should hold with such exactness in one domain of phenomena, and yet should be invalid for another, is *a priori* not very probable.

We now proceed to the second argument, to which, moreover, we shall return later. If the principle of relativity (in the restricted sense) does not hold, then the Galileian co-ordinate systems K, K', K'', etc., which are moving uniformly relative to each other, will not be *equivalent* for the description of natural phenomena. In this case we should be constrained to believe that natural laws are capable of being formulated in a particularly simple manner, and of course only on condition that, from amongst all possible Galileian co-ordinate systems, we should have chosen *one* (K_0) of a particular state of motion as our body of reference. We should then be justified (because of its merits for the description of natural phenomena) in calling this system "absolutely at rest," and all other Galileian systems K "in

motion." If, for instance, our embankment were the system K_0 , then our railway carriage would be a system K, relative to which less simple laws would hold than with respect to K_0 . This diminished simplicity would be due to the fact that the carriage K would be in motion (i.e. "really") with respect to K_0 . In the general laws of nature which have been formulated with reference to K, the magnitude and direction of the velocity of the carriage would necessarily play a part. We should expect, for instance, that the note emitted by an organ-pipe placed with its axis parallel to the direction of travel would be different from that emitted if the axis of the pipe were placed perpendicular to this direction. Now in virtue of its motion in an orbit round the sun, our earth is comparable with a railway carriage travelling with a velocity of about 30 kilometres per second. If the principle of relativity were not valid we should therefore expect that the direction of motion of the earth at any moment would enter into the laws of nature, and also that physical systems in their behaviour would be dependent on the orientation in space with respect to the earth. For owing to the alteration in direction of the velocity of revolution of the earth in the course of a year, the earth cannot be at rest relative to the hypothetical system K_0 throughout the whole year. However, the most careful observations have never revealed such anisotropic properties in terrestrial physical space, i.e. a physical nonequivalence of different directions. This is a very powerful argument in favour of the principle of relativity.

VI

THE THEOREM OF THE ADDITION OF VELOCITIES EMPLOYED IN CLASSICAL MECHANICS

ET us suppose our old friend the railway carriage to be λ travelling along the rails with a constant velocity ν , and that a man traverses the length of the carriage in the direction of travel with a velocity w. How quickly or, in other words, with what velocity W does the man advance relative to the embankment during the process? The only possible answer seems to result from the following consideration: If the man were to stand still for a second, he would advance relative to the embankment through a distance v equal numerically to the velocity of the carriage. As a consequence of his walking, however, he traverses an additional distance w relative to the carriage, and hence also relative to the embankment, in this second, the distance w being numerically equal to the velocity with which he is walking. Thus in total he covers the distance W = v + w relative to the embankment in the second considered. We shall see later that this result, which expresses the theorem of the addition of velocities employed in classical mechanics, cannot be maintained; in other words, the law that we have just written down does not hold in reality. For the time being, however, we shall assume its correctness.

VII

THE APPARENT INCOMPATIBILITY OF THE LAW OF PROPAGATION OF LIGHT WITH THE PRINCIPLE OF RELATIVITY

THERE is hardly a simpler law in physics than that according to which light is propagated in empty space. Every child at school knows, or believes he knows, that this propagation takes place in straight lines with a velocity $c = 300,000 \, \mathrm{km./sec.}$ At all events we know with great exactness that this velocity is the same for all colours, because if this were not the case, the minimum of emission would not be observed simultaneously for different colours during the eclipse of a fixed star by its dark neighbour. By means of similar considerations based on observations of double stars, the Dutch astronomer De Sitter was also able to show that the velocity of propagation of light cannot depend on the velocity of motion of the body emitting the light. The assumption that this velocity of propagation is dependent on the direction "in space" is in itself improbable.

In short, let us assume that the simple law of the constancy of the velocity of light c (in vacuum) is justifiably believed by the child at school. Who would imagine that this simple law has plunged the conscientiously thoughtful physicist into the greatest intellectual difficulties? Let us consider how these difficulties arise.

Of course we must refer the process of the propagation of light (and indeed every other process) to a rigid reference-body (coordinate system). As such a system let us again choose our embankment. We shall imagine the air above it to have been removed. If a ray of light be sent along the embankment, we see from the above that the tip of the ray will be transmitted with the velocity c relative to the embankment. Now let us suppose that our railway carriage is again travelling along the railway lines with the velocity v, and that its direction is the same as that of the ray of light, but its velocity of course much less. Let us inquire about the velocity of propagation of the ray of light relative to the carriage. It is obvious that we can here apply the consideration of the previous section, since the ray of light plays the part of the man walking along relatively to the carriage. The velocity W of the man relative to the embankment is here replaced by the velocity of light relative to the embankment. w is the required velocity of light with respect to the carriage, and we have

$$w = c - v$$
.

The velocity of propagation of a ray of light relative to the carriage thus comes out smaller than c.

But this result comes into conflict with the principle of relativity set forth in Section V. For, like every other general law of nature, the law of the transmission of light *in vacuo* must, according to the principle of relativity, be the same for the railway carriage as reference-body as when the rails are the body of reference. But, from our above consideration, this would appear to be impossible. If every ray of light is propagated relative to the embankment with the velocity c, then for this reason it would appear that another law of propagation of light must necessarily hold with respect to the carriage—a result contradictory to the principle of relativity.

In view of this dilemma there appears to be nothing else for it than to abandon either the principle of relativity or the simple law of the propagation of light *in vacuo*. Those of you who have carefully followed the preceding discussion are almost sure to expect that we should retain the principle of relativity, which appeals so convincingly to the intellect because it is so natural and simple. The law of the propagation of light *in vacuo* would then have to be replaced by a more complicated law conformable to the principle of relativity. The development of theoretical physics shows, however, that we cannot pursue this course. The epoch-making theoretical

investigations of H. A. Lorentz on the electrodynamical and optical phenomena connected with moving bodies show that experience in this domain leads conclusively to a theory of electromagnetic phenomena, of which the law of the constancy of the velocity of light *in vacuo* is a necessary consequence. Prominent theoretical physicists were therefore more inclined to reject the principle of relativity, in spite of the fact that no empirical data had been found which were contradictory to this principle.

At this juncture the theory of relativity entered the arena. As a result of an analysis of the physical conceptions of time and space, it became evident that *in reality there is not the least incompatibility between the principle of relativity and the law of propagation of light,* and that by systematically holding fast to both these laws a logically rigid theory could be arrived at. This theory has been called the *special theory of relativity* to distinguish it from the extended theory, with which we shall deal later. In the following pages we shall present the fundamental ideas of the special theory of relativity.

VIII

ON THE IDEA OF TIME IN PHYSICS

IGHTNING has struck the rails on our railway embankment at two places A and B far distant from each other. I make the additional assertion that these two lightning flashes occurred simultaneously. If I ask you whether there is sense in this statement, you will answer my question with a decided "Yes." But if I now approach you with the request to explain to me the sense of the statement more precisely, you find after some consideration that the answer to this question is not so easy as it appears at first sight.

After some time perhaps the following answer would occur to you: "The significance of the statement is clear in itself and needs no further explanation; of course it would require some consideration if I were to be commissioned to determine by observations whether in the actual case the two events took place simultaneously or not." I cannot be satisfied with this answer for the following reason. Supposing that as a result of ingenious considerations an able meteorologist were to discover that the lightning must always strike the places A and B simultaneously, then we should be faced with the task of testing whether or not this theoretical result is in accordance with the reality. We encounter the same difficulty with all physical statements in which the conception "simultaneous" plays a part. The concept does not exist for the physicist until he has the possibility of discovering whether or not it is fulfilled in an actual case. We thus require a definition of simultaneity such that this definition supplies us with the method by means of which, in the present case, he can decide by experiment whether or not both the lightning strokes occurred simultaneously. As long as this requirement is not satisfied, I allow

myself to be deceived as a physicist (and of course the same applies if I am not a physicist), when I imagine that I am able to attach a meaning to the statement of simultaneity. (I would ask the reader not to proceed farther until he is fully convinced on this point.)

After thinking the matter over for some time you then offer the following suggestion with which to test simultaneity. By measuring along the rails, the connecting line AB should be measured up and an observer placed at the mid-point M of the distance AB. This observer should be supplied with an arrangement (e.g. two mirrors inclined at 90°) which allows him visually to observe both places A and B at the same time. If the observer perceives the two flashes of lightning at the same time, then they are simultaneous.

I am very pleased with this suggestion, but for all that I cannot regard the matter as quite settled, because I feel constrained to raise the following objection: "Your definition would certainly be right, if I only knew that the light by means of which the observer at M perceives the lightning flashes travels along the length $A \longrightarrow M$. With the same velocity as along the length $B \longrightarrow M$. But an examination of this supposition would only be possible if we already had at our disposal the means of measuring time. It would thus appear as though we were moving here in a logical circle."

After further consideration you cast a somewhat disdainful glance at me—and rightly so—and you declare: "I maintain my previous definition nevertheless, because in reality it assumes absolutely nothing about light. There is only *one* demand to be made of the definition of simultaneity, namely, that in every real case it must supply us with an empirical decision as to whether or not the conception that has to be defined is fulfilled. That my definition satisfies this demand is indisputable. That light requires the same time to traverse the path $A \longrightarrow M$ as for the path $B \longrightarrow M$ is in reality neither a *supposition nor a hypothesis* about the physical nature of light, but a *stipulation* which I can make of my own freewill in order to arrive at a definition of simultaneity."

It is clear that this definition can be used to give an exact meaning not only to *two* events, but to as many events as we care to choose, and independently of the positions of the scenes of the events with

respect to the body of reference ¹ (here the railway embankment). We are thus led also to a definition of "time" in physics. For this purpose we suppose that clocks of identical construction are placed at the points *A*, *B* and *C* of the railway line (co-ordinate system), and that they are set in such a manner that the positions of their pointers are simultaneously (in the above sense) the same. Under these conditions we understand by the "time" of an event the reading (position of the hands) of that one of these clocks which is in the immediate vicinity (in space) of the event. In this manner a time-value is associated with every event which is essentially capable of observation.

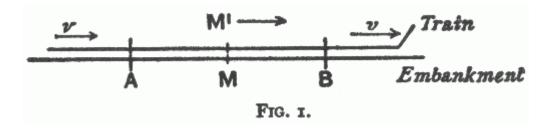
This stipulation contains a further physical hypothesis, the validity of which will hardly be doubted without empirical evidence to the contrary. It has been assumed that all these clocks go *at the same rate* if they are of identical construction. Stated more exactly: When two clocks arranged at rest in different places of a reference-body are set in such a manner that a *particular* position of the pointers of the one clock is *simultaneous* (in the above sense) with the *same* position of the pointers of the other clock, then identical "settings" are always simultaneous (in the sense of the above definition).

¹ We suppose further that, when three events A, B and C take place in different places in such a manner that, if A is simultaneous with B, and B is simultaneous with C (simultaneous in the sense of the above definition), then the criterion for the simultaneity of the pair of events A, C is also satisfied. This assumption is a physical hypothesis about the law of propagation of light; it must certainly be fulfilled if we are to maintain the law of the constancy of the velocity of light *in vacuo*.

IX

THE RELATIVITY OF SIMULTANEITY

P to now our considerations have been referred to a particular body of reference, which we have styled a "railway embankment." We suppose a very long train travelling along the rails with the constant velocity v and in the direction indicated in Fig. 1. People travelling in this train will with advantage use the train as a rigid reference-body (co-ordinate system); they regard all events



in reference to the train. Then every event which takes place along the line also takes place at a particular point of the train. Also the definition of simultaneity can be given relative to the train in exactly the same way as with respect to the embankment. As a natural consequence, however, the following question arises:

Are two events (e.g. the two strokes of lightning A and B) which are simultaneous with reference to the railway embankment also simultaneous relatively to the train? We shall show directly that the answer must be in the negative.

When we say that the lightning strokes A and B are simultaneous with respect to the embankment, we mean: the rays of light emitted at the places A and B, where the lightning occurs, meet each other at the mid-point M of the length $A \longrightarrow B$ of the embankment. But the

events A and B also correspond to positions A and B on the train. Let M' be the mid-point of the distance $A \longrightarrow B$ on the travelling train. Just when the flashes 1 of lightning occur, this point M' naturally coincides with the point M, but it moves towards the right in the diagram with the velocity v of the train. If an observer sitting in the position M' in the train did not possess this velocity, then he would remain permanently at M, and the light rays emitted by the flashes of lightning A and B would reach him simultaneously, i.e. they would meet just where he is situated. Now in reality (considered with reference to the railway embankment) he is hastening towards the beam of light coming from B, whilst he is riding on ahead of the beam of light coming from A. Hence the observer will see the beam of light emitted from B earlier than he will see that emitted from A. Observers who take the railway train as their reference-body must therefore come to the conclusion that the lightning flash B took place earlier than the lightning flash A. We thus arrive at the important result:

Events which are simultaneous with reference to the embankment are not simultaneous with respect to the train, and *vice versa* (relativity of simultaneity). Every reference-body (co-ordinate system) has its own particular time; unless we are told the reference-body to which the statement of time refers, there is no meaning in a statement of the time of an event.

Now before the advent of the theory of relativity it had always tacitly been assumed in physics that the statement of time had an absolute significance, *i.e.* that it is independent of the state of motion of the body of reference. But we have just seen that this assumption is incompatible with the most natural definition of simultaneity; if we discard this assumption, then the conflict between the law of the propagation of light *in vacuo* and the principle of relativity (developed in Section VII) disappears.

We were led to that conflict by the considerations of Section VI, which are now no longer tenable. In that section we concluded that the man in the carriage, who traverses the distance w per second

¹ As judged from the embankment.

relative to the carriage, traverses the same distance also with respect to the embankment *in each second* of time. But, according to the foregoing considerations, the time required by a particular occurrence with respect to the carriage must not be considered equal to the duration of the same occurrence as judged from the embankment (as reference-body). Hence it cannot be contended that the man in walking travels the distance *w* relative to the railway line in a time which is equal to one second as judged from the embankment.

Moreover, the considerations of Section VI are based on yet a second assumption, which, in the light of a strict consideration, appears to be arbitrary, although it was always tacitly made even before the introduction of the theory of relativity.

X

ON THE RELATIVITY OF THE CONCEPTION OF DISTANCE

ET us consider two particular points on the train ¹ travelling along the embankment with the velocity v, and inquire as to their distance apart. We already know that it is necessary to have a body of reference for the measurement of a distance, with respect to which body the distance can be measured up. It is the simplest plan to use the train itself as the reference-body (co-ordinate system). An observer in the train measures the interval by marking off his measuring-rod in a straight line (e.g. along the floor of the carriage) as many times as is necessary to take him from the one marked point to the other. Then the number which tells us how often the rod has to be laid down is the required distance.

It is a different matter when the distance has to be judged from the railway line. Here the following method suggests itself. If we call A' and B' the two points on the train whose distance apart is required, then both of these points are moving with the velocity v along the embankment. In the first place we require to determine the points A and B of the embankment which are just being passed by the two points A' and B' at a particular time t—judged from the embankment. These points A and B of the embankment can be determined by applying the definition of time given in Section VIII. The distance between these points A and B is then measured by repeated application of the measuring-rod along the embankment.

¹ e.g. the middle of the first and of the hundredth carriage.

A priori it is by no means certain that this last measurement will supply us with the same result as the first. Thus the length of the train as measured from the embankment may be different from that obtained by measuring in the train itself. This circumstance leads us to a second objection which must be raised against the apparently obvious consideration of Section VI. Namely, if the man in the carriage covers the distance w in a unit of time—measured from the train,—then this distance—as measured from the embankment—is not necessarily also equal to w.

XI

THE LORENTZ TRANSFORMATION

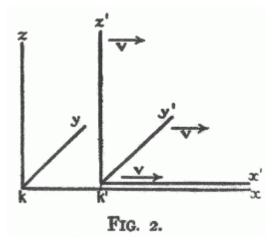
THE results of the last three sections show that the apparent incompatibility of the law of propagation of light with the principle of relativity (Section VII) has been derived by means of a consideration which borrowed two unjustifiable hypotheses from classical mechanics; these are as follows:

- (1) The time-interval (time) between two events is independent of the condition of motion of the body of reference.
- (2) The space-interval (distance) between two points of a rigid body is independent of the condition of motion of the body of reference.

If we drop these hypotheses, then the dilemma of Section VII disappears, because the theorem of the addition of velocities derived in Section VI becomes invalid. The possibility presents itself that the law of the propagation of light in vacuo may be compatible with the principle of relativity, and the question arises: How have we to modify the considerations of Section VI in order to remove the apparent disagreement between these two fundamental results of experience? This question leads to a general one. In the discussion of Section VI we have to do with places and times relative both to the train and to the embankment. How are we to find the place and time of an event in relation to the train, when we know the place and time of the event with respect to the railway embankment? Is there a thinkable answer to this question of such a nature that the law of transmission of light in vacuo does not contradict the principle of relativity? In other words: Can we conceive of a relation between place and time of the individual events relative to both referencebodies, such that every ray of light possesses the velocity of transmission c relative to the embankment and relative to the train? This question leads to a quite definite positive answer, and to a perfectly definite transformation law for the space-time magnitudes of an event when changing over from one body of reference to another.

Before we deal with this, we shall introduce the following incidental consideration. Up to the present we have only considered events taking place along the embankment, which had mathematically to assume the function of a straight line. In the manner indicated in Section II we can imagine this reference-body supplemented laterally and in a vertical direction by means of a framework of rods, so that an event which takes place anywhere can be localised with reference to this framework. Similarly, we can imagine the train travelling with the velocity v to be continued across the whole of space, so that every event, no matter how far off it may be, could also be localised with second framework. Without committing the fundamental error, we can disregard the fact that in reality these frameworks would continually interfere with each other, owing to the impenetrability of solid bodies. In every such framework we imagine three surfaces perpendicular to each other marked out, and designated as "co-ordinate planes" ("co-ordinate system"). A co-ordinate system K then corresponds to the embankment, and a co-ordinate system K'to the train. An event, wherever it may have taken place, would be fixed in space with respect to K by the three perpendiculars x, y, z on

the co-ordinate planes, and with regard to time by a time-value t. Relative to K', the same event would be fixed in respect of space and time by corresponding values x', y', z', t', which of course are not identical with x, y, z, t. It has already been set forth in detail how these magnitudes are to be regarded as results of physical measurements.



Obviously our problem can be exactly formulated in the following manner. What are the values x', y', z', t' of an event with respect to K', when the magnitudes x, y, z, t, of the same event with respect to K are

given? The relations must be so chosen that the law of the transmission of light *in vacuo* is satisfied for one and the same ray of light (and of course for every ray) with respect to K and K'. For the relative orientation in space of the co-ordinate systems indicated in the diagram (Fig. 2), this problem is solved by means of the equations:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$y' = y$$

$$z' = z$$

$$t' = \frac{t - \frac{v}{c^2} \cdot x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

This system of equations is known as the "Lorentz transformation." ¹

If in place of the law of transmission of light we had taken as our basis the tacit assumptions of the older mechanics as to the absolute character of times and lengths, then instead of the above we should have obtained the following equations:

$$x' = x - vt$$

$$y' = y$$

$$z' = z$$

$$t' = t$$

This system of equations is often termed the "Galilei transformation." The Galilei transformation can be obtained from the Lorentz transformation by substituting an infinitely large value for the velocity of light c in the latter transformation.

¹ A simple derivation of the Lorentz transformation is given in Appendix I.

Aided by the following illustration, we can readily see that, in accordance with the Lorentz transformation, the law of the transmission of light *in vacuo* is satisfied both for the reference-body K and for the reference-body K'. A light-signal is sent along the positive x-axis, and this light-stimulus advances in accordance with the equation

$$x = ct$$

i.e. with the velocity c. According to the equations of the Lorentz transformation, this simple relation between x and t involves a relation between x' and t'. In point of fact, if we substitute for x the value ct in the first and fourth equations of the Lorentz transformation, we obtain:

$$x' = \frac{(c-v)t}{\sqrt{1-\frac{v^2}{c^2}}}$$
$$t' = \frac{\left(1-\frac{v}{c}\right)t}{\sqrt{1-\frac{v^2}{c^2}}},$$

from which, by division, the expression

$$x' = ct'$$

immediately follows. If referred to the system K', the propagation of light takes place according to this equation. We thus see that the velocity of transmission relative to the reference-body K' is also equal to c. The same result is obtained for rays of light advancing in any other direction whatsoever. Of course this is not surprising, since the equations of the Lorentz transformation were derived conformably to this point of view.

XII

THE BEHAVIOUR OF MEASURING-RODS AND CLOCKS IN MOTION

PLACE a metre-rod in the x'-axis of K' in such a manner that one end (the beginning) coincides with the point x' = 0, whilst the other end (the end of the rod) coincides with the point x' = 1. What is the length of the metre-rod relatively to the system K? In order to learn this, we need only ask where the beginning of the rod and the end of the rod lie with respect to K at a particular time t of the system K. By means of the first equation of the Lorentz transformation the values of these two points at the time t = 0 can be shown to be

$$x_{\text{(beginning of rod)}} = 0 \cdot \sqrt{1 - \frac{v^2}{c^2}}$$
$$x_{\text{(end of rod)}} = 1 \cdot \sqrt{1 - \frac{v^2}{c^2}},$$

the distance between the points being $\sqrt{1-\frac{v^2}{c^2}}$. But the metre-rod is moving with the velocity v relative to K. It therefore follows that the length of a rigid metre-rod moving in the direction of its length with a velocity v is $\sqrt{1-v^2/c^2}$ of a metre. The rigid rod is thus shorter when in motion than when at rest, and the more quickly it is moving, the shorter is the rod. For the velocity v=c we should have

 $\sqrt{1-v^2/c^2} = 0$, and for still greater velocities the square-root becomes imaginary. From this we conclude that in the theory of relativity the velocity c plays the part of a limiting velocity, which can neither be reached nor exceeded by any real body.

Of course this feature of the velocity c as a limiting velocity also clearly follows from the equations of the Lorentz transformation, for these become meaningless if we choose values of v greater than c.

If, on the contrary, we had considered a metre-rod at rest in the x-axis with respect to K, then we should have found that the length of the rod as judged from K' would have been $\sqrt{1-v^2/c^2}$; this is quite in accordance with the principle of relativity which forms the basis of our considerations.

A priori it is quite clear that we must be able to learn something about the physical behaviour of measuring-rods and clocks from the equations of transformation, for the magnitudes x, y, z, t, are nothing more nor less than the results of measurements obtainable by means of measuring-rods and clocks. If we had based our considerations on the Galilei transformation we should not have obtained a contraction of the rod as a consequence of its motion.

Let us now consider a seconds-clock which is permanently situated at the origin (x'=0) of K'. t'=0 and t'=1 are two successive ticks of this clock. The first and fourth equations of the Lorentz transformation give for these two ticks:

$$t = 0$$

and

$$t = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

As judged from K, the clock is moving with the velocity v; as judged from this reference-body, the time which elapses between two

strokes of the clock is not one second, but $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ seconds, *i.e.* a

somewhat larger time. As a consequence of its motion the clock goes more slowly than when at rest. Here also the velocity c plays the part of an unattainable limiting velocity.

XIII

THEOREM OF THE ADDITION OF VELOCITIES. THE EXPERIMENT OF FIZEAU

OW in practice we can move clocks and measuring-rods only with velocities that are small compared with the velocity of light; hence we shall hardly be able to compare the results of the previous section directly with the reality. But, on the other hand, these results must strike you as being very singular, and for that reason I shall now draw another conclusion from the theory, one which can easily be derived from the foregoing considerations, and which has been most elegantly confirmed by experiment.

In Section VI we derived the theorem of the addition of velocities in one direction in the form which also results from the hypotheses of classical mechanics. This theorem can also be deduced readily from the Galilei transformation (Section XI). In place of the man walking inside the carriage, we introduce a point moving relatively to the coordinate system K' in accordance with the equation

$$x' = wt'$$
.

By means of the first and fourth equations of the Galilei transformation we can express x' and t' in terms of x and t, and we then obtain

$$x = (v + w)t.$$

This equation expresses nothing else than the law of motion of the point with reference to the system K (of the man with reference to the

embankment). We denote this velocity by the symbol W, and we then obtain, as in Section VI,

$$W = v + w$$
 (A).

But we can carry out this consideration just as well on the basis of the theory of relativity. In the equation

$$x' = wt'$$

we must then express x' and t' in terms of x and t, making use of the first and fourth equations of the *Lorentz transformation*. Instead of the equation (A) we then obtain the equation

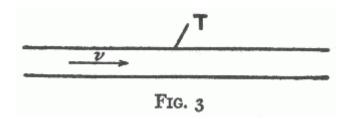
$$W = \frac{v+w}{1+\frac{vw}{c^2}} \qquad . \qquad . \qquad . \qquad . \qquad . \qquad (B),$$

which corresponds to the theorem of addition for velocities in one direction according to the theory of relativity. The question now arises as to which of these two theorems is the better in accord with experience. On this point we are enlightened by a most important experiment which the brilliant physicist Fizeau performed more than half a century ago, and which has been repeated since then by some of the best experimental physicists, so that there can be no doubt about its result. The experiment is concerned with the following question. Light travels in a motionless liquid with a particular velocity w. How quickly does it travel in the direction of the arrow in the tube T (see the accompanying diagram, Fig. 3) when the liquid above mentioned is flowing through the tube with a velocity v?

In accordance with the principle of relativity we shall certainly have to take for granted that the propagation of light always takes place with the same velocity *w* with respect to the liquid, whether the latter is in motion with reference to other bodies or not. The velocity

of light relative to the liquid and the velocity of the latter relative to the tube are thus known, and we require the velocity of light relative to the tube.

It is clear that we have the problem of Section VI again before us. The tube plays the part of the railway embankment or of the co-



ordinate system K, the liquid plays the part of the carriage or of the co-ordinate system K', and finally, the light plays the part of the man walking along the carriage, or of the moving point in the present section. If we denote the velocity of the light relative to the tube by W, then this is given by the equation (A) or (B), according as the Galilei transformation or the Lorentz transformation corresponds to the facts. Experiment ¹ decides in favour of equation (B) derived from the theory of relativity, and the agreement is, indeed, very exact. According to recent and most excellent measurements by Zeeman, the influence of the velocity of flow v on the propagation of light is represented by formula (B) to within one per cent.

Nevertheless we must now draw attention to the fact that a theory of this phenomenon was given by H. A. Lorentz long before the statement of the theory of relativity. This theory was of a purely electrodynamical nature, and was obtained by the use of particular hypotheses as to the electromagnetic structure of matter. This circumstance, however, does not in the least diminish the conclusiveness of the experiment as a crucial test in favour of the

¹ Fizeau found $W = w + v \left(1 - \frac{1}{n^2}\right)$, where $n = \frac{c}{w}$ is the index of refraction of the liquid. On the other hand, owing to the smallness of $\frac{vw}{c^2}$ as compared with 1, we can replace (B) in the first place by $W = (w + v) \left(1 - \frac{vw}{c^2}\right)$, or to the same order of approximation by $w + v \left(1 - \frac{1}{n^2}\right)$, which agrees with Fizeau's result.

theory of relativity, for the electrodynamics of Maxwell-Lorentz, on which the original theory was based, in no way opposes the theory of relativity. Rather has the latter been developed from electrodynamics as an astoundingly simple combination and generalisation of the hypotheses, formerly independent of each other, on which electrodynamics was built.

XIV

THE HEURISTIC VALUE OF THE THEORY OF RELATIVITY

UR train of thought in the foregoing pages can be epitomised in the following manner. Experience has led to the conviction that, on the one hand, the principle of relativity holds true, and that on the other hand the velocity of transmission of light *in vacuo* has to be considered equal to a constant *c*. By uniting these two postulates we obtained the law of transformation for the rectangular co-ordinates *x*, *y*, *z* and the time *t* of the events which constitute the processes of nature. In this connection we did not obtain the Galilei transformation, but, differing from classical mechanics, the *Lorentz transformation*.

The law of transmission of light, the acceptance of which is justified by our actual knowledge, played an important part in this process of thought. Once in possession of the Lorentz transformation, however, we can combine this with the principle of relativity, and sum up the theory thus:

Every general law of nature must be so constituted that it is transformed into a law of exactly the same form when, instead of the space-time variables x, y, z, t of the original co-ordinate system K, we introduce new space-time variables x', y', z', t' of a co-ordinate system K'. In this connection the relation between the ordinary and the accented magnitudes is given by the Lorentz transformation. Or in brief: General laws of nature are co-variant with respect to Lorentz transformations.

This is a definite mathematical condition that the theory of relativity demands of a natural law, and in virtue of this, the theory becomes a valuable heuristic aid in the search for general laws of nature. If a general law of nature were to be found which did not satisfy this condition, then at least one of the two fundamental assumptions of the theory would have been disproved. Let us now examine what general results the latter theory has hitherto evinced.

XV

GENERAL RESULTS OF THE THEORY

T is clear from our previous considerations that the (special) theory of relativity has grown out of electrodynamics and optics. In these fields it has not appreciably altered the predictions of theory, but it has considerably simplified the theoretical structure, *i.e.* the derivation of laws, and—what is incomparably more important—it has considerably reduced the number of independent hypotheses forming the basis of theory. The special theory of relativity has rendered the Maxwell-Lorentz theory so plausible, that the latter would have been generally accepted by physicists even if experiment had decided less unequivocally in its favour.

Classical mechanics required to be modified before it could come into line with the demands of the special theory of relativity. For the main part, however, this modification affects only the laws for rapid motions, in which the velocities of matter ν are not very small as compared with the velocity of light. We have experience of such rapid motions only in the case of electrons and ions; for other motions the variations from the laws of classical mechanics are too small to make themselves evident in practice. We shall not consider the motion of stars until we come to speak of the general theory of relativity. In accordance with the theory of relativity the kinetic energy of a material point of mass m is no longer given by the well-known expression

$$m\frac{v^2}{2}$$
,

but by the expression

$$\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}.$$

This expression approaches infinity as the velocity v approaches the velocity of light c. The velocity must therefore always remain less than c, however great may be the energies used to produce the acceleration. If we develop the expression for the kinetic energy in the form of a series, we obtain

$$mc^2 + m\frac{v^2}{2} + \frac{3}{8}m\frac{v^4}{c^2} + \dots$$

When $\frac{v^2}{c^2}$ is small compared with unity, the third of these terms is always small in comparison with the second, which last is alone considered in classical mechanics. The first term mc^2 does not contain the velocity, and requires no consideration if we are only dealing with the question as to how the energy of a point-mass depends on the velocity. We shall speak of its essential significance later.

The most important result of a general character to which the special theory of relativity has led is concerned with the conception of mass. Before the advent of relativity, physics recognised two conservation laws of fundamental importance, namely, the law of the conservation of energy and the law of the conservation of mass; these two fundamental laws appeared to be quite independent of each other. By means of the theory of relativity they have been united into one law. We shall now briefly consider how this unification came about, and what meaning is to be attached to it.

The principle of relativity requires that the law of the conservation of energy should hold not only with reference to a co-ordinate system K, but also with respect to every co-ordinate system K' which is in a

state of uniform motion of translation relative to K, or, briefly, relative to every "Galileian" system of co-ordinates. In contrast to classical mechanics, the Lorentz transformation is the deciding factor in the transition from one such system to another.

By means of comparatively simple considerations we are led to draw the following conclusion from these premises, in conjunction with the fundamental equations of the electrodynamics of Maxwell: A body moving with the velocity v, which absorbs 1 an amount of energy E_0 in the form of radiation without suffering an alteration in velocity in the process, has, as a consequence, its energy increased by an amount

$$\frac{E_0}{\sqrt{1-\frac{v^2}{c^2}}}.$$

In consideration of the expression given above for the kinetic energy of the body, the required energy of the body comes out to be

$$\frac{\left(m+\frac{E_0}{c^2}\right)c^2}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Thus the body has the same energy as a body of mass $\left(m + \frac{E_0}{c^2}\right)$ moving with the velocity v. Hence we can say: If a body takes up an amount of energy E_0 , then its inertial mass increases by an amount $\frac{E_0}{c^2}$; the inertial mass of a body is not a constant, but varies according to the change in the energy of the body. The inertial mass of a system of bodies can even be regarded as a measure of its energy. The law of

 $^{^{1}}$ $E_{\scriptscriptstyle 0}$ is the energy taken up, as judged from a co-ordinate system moving with the body.

the conservation of the mass of a system becomes identical with the law of the conservation of energy, and is only valid provided that the system neither takes up nor sends out energy. Writing the expression for the energy in the form

$$\frac{mc^{2}+E_{0}}{\sqrt{1-\frac{v^{2}}{c^{2}}}},$$

we see that the term mc^2 , which has hitherto attracted our attention, is nothing else than the energy possessed by the body ¹ before it absorbed the energy E_0 .

A direct comparison of this relation with experiment is not possible at the present time, owing to the fact that the changes in energy E_0 to which we can subject a system are not large enough to make themselves perceptible as a change in the inertial mass of the system. $\frac{E_0}{c^2}$ is too small in comparison with the mass m, which was present before the alteration of the energy. It is owing to this circumstance that classical mechanics was able to establish successfully the conservation of mass as a law of independent validity.

Let me add a final remark of a fundamental nature. The success of the Faraday-Maxwell interpretation of electromagnetic action at a distance resulted in physicists becoming convinced that there are no such things as instantaneous actions at a distance (not involving an intermediary medium) of the type of Newton's law of gravitation. According to the theory of relativity, action at a distance with the velocity of light always takes the place of instantaneous action at a distance or of action at a distance with an infinite velocity of transmission. This is connected with the fact that the velocity c plays a fundamental rôle in this theory. In Part II we shall see in what way this result becomes modified in the general theory of relativity.

¹ As judged from a co-ordinate system moving with the body.

XVI

EXPERIENCE AND THE SPECIAL THEORY OF RELATIVITY

O what extent is the special theory of relativity supported by experience? This question is not as: experience? This question is not easily answered for the reason already mentioned in connection with the fundamental experiment of Fizeau. The special theory of relativity has crystallised out from the Maxwell-Lorentz theory of electromagnetic phenomena. Thus all facts of experience which support the electromagnetic theory also support the theory of relativity. As being of particular importance, I mention here the fact that the theory of relativity enables us to predict the effects produced on the light reaching us from the fixed stars. These results are obtained in an exceedingly simple manner, and the effects indicated, which are due to the relative motion of the earth with reference to those fixed stars, are found to be in accord with experience. We refer to the yearly movement of the apparent position of the fixed stars resulting from the motion of the earth round the sun (aberration), and to the influence of the radial components of the relative motions of the fixed stars with respect to the earth on the colour of the light reaching us from them. The latter effect manifests itself in a slight displacement of the spectral lines of the light transmitted to us from a fixed star, as compared with the position of the same spectral lines when they are produced by a terrestrial source of light (Doppler principle). The experimental arguments in favour of the Maxwell-Lorentz theory, which are at the same time arguments in favour of the theory of relativity, are too numerous to be set forth here. In reality they limit the theoretical possibilities to such an extent, that no other theory than that of Maxwell and Lorentz has been able to hold its own when tested by experience.

But there are two classes of experimental facts hitherto obtained which can be represented in the Maxwell-Lorentz theory only by the introduction of an auxiliary hypothesis, which in itself—*i.e.* without making use of the theory of relativity—appears extraneous.

It is known that cathode rays and the so-called β -rays emitted by radioactive substances consist of negatively electrified particles (electrons) of very small inertia and large velocity. By examining the deflection of these rays under the influence of electric and magnetic fields, we can study the law of motion of these particles very exactly.

In the theoretical treatment of these electrons, we are faced with the difficulty that electrodynamic theory of itself is unable to give an account of their nature. For since electrical masses of one sign repel each other, the negative electrical masses constituting the electron would necessarily be scattered under the influence of their mutual repulsions, unless there are forces of another kind operating between them, the nature of which has hitherto remained obscure to us. If we now assume that the relative distances between the electrical masses constituting the electron remain unchanged during the motion of the electron (rigid connection in the sense of classical mechanics), we arrive at a law of motion of the electron which does not agree with experience. Guided by purely formal points of view, H. A. Lorentz was the first to introduce the hypothesis that the particles constituting the electron experience a contraction in the direction of motion in consequence of that motion, the amount of this contraction being

proportional to the expression $\sqrt{1-\frac{v^2}{c^2}}$.* This hypothesis, which is not justifiable by any electrodynamical facts, supplies us then with that particular law of motion which has been confirmed with great precision in recent years.

[* This expression was misprinted
$$\sqrt{1 = \frac{v^2}{c^2}}$$
 in the original book.— J.M.]

¹ The general theory of relativity renders it likely that the electrical masses of an electron are held together by gravitational forces.

The theory of relativity leads to the same law of motion, without requiring any special hypothesis whatsoever as to the structure and the behaviour of the electron. We arrived at a similar conclusion in Section XIII in connection with the experiment of Fizeau, the result of which is foretold by the theory of relativity without the necessity of drawing on hypotheses as to the physical nature of the liquid.

The second class of facts to which we have alluded has reference to the question whether or not the motion of the earth in space can be made perceptible in terrestrial experiments. We have already remarked in Section V that all attempts of this nature led to a negative result. Before the theory of relativity was put forward, it was difficult to become reconciled to this negative result, for reasons now to be discussed. The inherited prejudices about time and space did not allow any doubt to arise as to the prime importance of the Galilei transformation for changing over from one body of reference to another. Now assuming that the Maxwell-Lorentz equations hold for a reference-body K, we then find that they do not hold for a referencebody K' moving uniformly with respect to K, if we assume that the relations of the Galileian transformation exist between the coordinates of K and K'. It thus appears that of all Galileian co-ordinate systems one (K) corresponding to a particular state of motion is physically unique. This result was interpreted physically by regarding K as at rest with respect to a hypothetical æther of space. On the other hand, all co-ordinate systems K' moving relatively to K were to be regarded as in motion with respect to the æther. To this motion of K'against the α ther (" α ther-drift" relative to K) were assigned the more complicated laws which were supposed to hold relative to K'. Strictly speaking, such an æther-drift ought also to be assumed relative to the earth, and for a long time the efforts of physicists were devoted to attempts to detect the existence of an æther-drift at the earth's surface.

In one of the most notable of these attempts Michelson devised a method which appears as though it must be decisive. Imagine two mirrors so arranged on a rigid body that the reflecting surfaces face each other. A ray of light requires a perfectly definite time T to pass from one mirror to the other and back again, if the whole system be at rest with respect to the æther. It is found by calculation, however, that

a slightly different time T' is required for this process, if the body, together with the mirrors, be moving relatively to the æther. And yet another point: it is shown by calculation that for a given velocity v with reference to the α ther, this time T' is different when the body is moving perpendicularly to the planes of the mirrors from that resulting when the motion is parallel to these planes. Although the estimated difference between these two times is exceedingly small, and Morley performed an experiment involving interference in which this difference should have been clearly detectable. But the experiment gave a negative result—a fact very perplexing to physicists. Lorentz and FitzGerald rescued the theory from this difficulty by assuming that the motion of the body relative to the æther produces a contraction of the body in the direction of motion, the amount of contraction being just sufficient to compensate for the difference in time mentioned above. Comparison with the discussion in Section XII shows that also from the standpoint of the theory of relativity this solution of the difficulty was the right one. But on the basis of the theory of relativity the method of interpretation is incomparably more satisfactory. According to this theory there is no such thing as a "specially favoured" (unique) coordinate system to occasion the introduction of the æther-idea, and hence there can be no æther-drift, nor any experiment with which to demonstrate it. Here the contraction of moving bodies follows from the two fundamental principles of the theory without the introduction of particular hypotheses; and as the prime factor involved in this contraction we find, not the motion in itself, to which we cannot attach any meaning, but the motion with respect to the body of reference chosen in the particular case in point. Thus for a co-ordinate system moving with the earth the mirror system of Michelson and Morley is not shortened, but it is shortened for a co-ordinate system which is at rest relatively to the sun.

XVII

MINKOWSKI'S FOUR-DIMENSIONAL SPACE

THE non-mathematician is seized by a mysterious shuddering when he hears of "four-dimensional" things, by a feeling not unlike that awakened by thoughts of the occult. And yet there is no more common-place statement than that the world in which we live is a four-dimensional space-time continuum.

Space is a three-dimensional continuum. By this we mean that it is possible to describe the position of a point (at rest) by means of three numbers (co-ordinates) x, y, z, and that there is an indefinite number of points in the neighbourhood of this one, the position of which can be described by co-ordinates such as x_1 , y_1 , z_1 , which may be as near as we choose to the respective values of the co-ordinates x, y, z of the first point. In virtue of the latter property we speak of a "continuum," and owing to the fact that there are three co-ordinates we speak of it as being "three-dimensional."

Similarly, the world of physical phenomena which was briefly called "world" by Minkowski is naturally four-dimensional in the space-time sense. For it is composed of individual events, each of which is described by four numbers, namely, three space co-ordinates x, y, z and a time co-ordinate, the time-value t. The "world" is in this sense also a continuum; for to every event there are as many "neighbouring" events (realised or at least thinkable) as we care to choose, the co-ordinates x_1 , y_1 , z_1 , t_1 of which differ by an indefinitely small amount from those of the event x, y, z, t originally considered. That we have not been accustomed to regard the world in this sense as a four-dimensional continuum is due to the fact that in physics, before the advent of the theory of relativity, time played a different

and more independent rôle, as compared with the space co-ordinates. It is for this reason that we have been in the habit of treating time as an independent continuum. As a matter of fact, according to classical mechanics, time is absolute, *i.e.* it is independent of the position and the condition of motion of the system of co-ordinates. We see this expressed in the last equation of the Galileian transformation (t'=t).

The four-dimensional mode of consideration of the "world" is natural on the theory of relativity, since according to this theory time is robbed of its independence. This is shown by the fourth equation of the Lorentz transformation:

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Moreover, according to this equation the time difference $\Delta t'$ of two events with respect to K' does not in general vanish, even when the time difference Δt of the same events with reference to K vanishes. Pure "space-distance" of two events with respect to K results in "time-distance" of the same events with respect to K'. But the discovery of Minkowski, which was of importance for the formal development of the theory of relativity, does not lie here. It is to be found rather in the fact of his recognition that the four-dimensional space-time continuum of the theory of relativity, in its most essential formal properties, shows a pronounced relationship to the threedimensional continuum of Euclidean geometrical space. In order to give due prominence to this relationship, however, we must replace the usual time co-ordinate t by an imaginary magnitude $\sqrt{-1 \cdot ct}$ proportional to it. Under these conditions, the natural laws satisfying the demands of the (special) theory of relativity assume mathematical forms, in which the time co-ordinate plays exactly the same rôle as the three space co-ordinates. Formally, these four co-ordinates

¹ Cf. the somewhat more detailed discussion in Appendix II.

correspond exactly to the three space co-ordinates in Euclidean geometry. It must be clear even to the non-mathematician that, as a consequence of this purely formal addition to our knowledge, the theory perforce gained clearness in no mean measure.

These inadequate remarks can give the reader only a vague notion of the important idea contributed by Minkowski. Without it the general theory of relativity, of which the fundamental ideas are developed in the following pages, would perhaps have got no farther than its long clothes. Minkowski's work is doubtless difficult of access to anyone inexperienced in mathematics, but since it is not necessary to have a very exact grasp of this work in order to understand the fundamental ideas of either the special or the general theory of relativity, I shall at present leave it here, and shall revert to it only towards the end of Part II.