Ether and the Theory of Relativity

Albert Einstein, an address delivered on May 5th, 1920, in the University of Leyden.

The original version is available in the Collected Papers of Albert Einstein.

See also the Einstein Archives Online.

HOW does it come about that alongside of the idea of ponderable matter, which is derived by abstraction from everyday life, the physicists set the idea of the existence of another kind of matter, the ether? The explanation is probably to be sought in those phenomena which have given rise to the theory of action at a distance, and in the properties of light which have led to the undulatory theory. Let us devote a little while to the consideration of these two subjects.

Outside of physics we know nothing of action at a distance. When we try to connect cause and effect in the experiences which natural objects afford us, it seems at first as if there were no other mutual actions than those of immediate contact, e.g. the communication of motion by impact, push and pull, heating or inducing combustion by means of a flame, etc. It is true that even in everyday experience weight, which is in a sense action at a distance, plays a very important part. But since in daily experience the weight of bodies meets us as something constant, something not linked to any cause which is variable in time or place, we do not in everyday life speculate as to the cause of gravity, and therefore do not become conscious of its character as action at a distance. It was Newton's theory of gravitation that first assigned a cause for gravity by interpreting it as action at a distance, proceeding from masses. Newton's theory is probably the greatest stride ever made in the effort towards the causal nexus of natural phenomena. And yet this theory evoked a lively sense of discomfort among Newton's contemporaries, because it seemed to be in conflict with the principle springing from the rest of experience, that there can be reciprocal action only through contact, and not through immediate action at a distance.

It is only with reluctance that man's desire for knowledge endures a dualism of thls kind. How was unity to be preserved in his comprehension of the forces of nature? Either by trying to look upon contact forces as being themselves distant forces which admittedly are observable only at a very small distance and this was the road which Newton's followers, who were entirely under the spell of his doctrine, mostly preferred to take; or by assuming that the Newtonian action at a distance is only apparently immediate action at a distance, but in truth is conveyed by a medium permeating space, whether by movements or by elastic deformation of this medium. Thus the endeavour toward a unified view of the nature of forces leads to the hypothesis of an ether. This hypothesis, to be sure, did not at first bring with it any advance in the theory of gravitation or in physics generally, so that it became customary to treat Newton's law of force as an axiom not further reducible. But the ether hypothesis was bound always to play some part in physical science, even if at first only a latent part.

When in the first half of the nineteenth century the far-reaching similarity was revealed which subsists between the properties of light and those of elastic waves in ponderable bodies, the ether hypothesis found fresh support. 1t appeared beyond question that light must be interpreted as a vibratory process in an elastic, inert medium filling up universal space. It also seemed to be a necessary consequence of the fact that light is capable of polarisation that this medium, the ether,

must be of the nature of a solid body, because transverse waves are not possible in a fluid, but only in a solid. Thus the physicists were bound to arrive at the theory of the ``quas-irigid'' luminiferous ether, the parts of which can carry out no movements relatively to one another except the small movements of deformation which correspond to light-waves.

This theory also called the theory of the stationary luminiferous ether moreover found a strong support in an experiment which is also of fundamental importance in the special theory of relativity, the experiment of Fizeau, from which one was obliged to infer that the luminiferous ether does not take part in the movements of bodies. The phenomenon of aberration also favoured the theory of the quasi-rigid ether.

The development of the theory of electricity along the path opened up by Maxwell and Lorentz gave the development of our ideas concerning the ether quite a peculiar and unexpected turn. For Maxwell himself the ether indeed still had properties which were purely mechanical, although of a much more complicated kind than the mechanical properties of tangible solid bodies. But neither Maxwell nor his followers succeeded in elaborating a mechanical model for the ether which might furnish a satisfactory mechanical interpretation of Maxwell's laws of the electro-magnetic field. The laws were clear and simple, the mechanical interpretations clumsy and contradictory. Almost imperceptibly the theoretical physicists adapted themselves to a situation which, from the standpoint of their mechanical programme, was very depressing. They were particularly influenced by the electro-dynamical investigations of Heinrich Hertz. For whereas they previously had required of a conclusive theory that it should content itself with the fundamental concepts which belong exclusively to mechanics (e.g. densities, velocities, deformations, stresses) they gradually accustomed themselves to admitting electric and magnetic force as fundamental concepts side by side with those of mechanics, without requiring a mechanical interpretation for them. Thus the purely mechanical view of nature was gradually abandoned. But this change led to a fundamental dualism which in the long-run was insupportable. A way of escape was now sought in the reverse direction, by reducing the principles of mechanics to those of electricity, and this especially as confidence in the strict validity of the equations of Newton's mechanics was shaken by the experiments with b-rays and rapid kathode rays.

This dualism still confronts us in unextenuated form in the theory of Hertz, where matter appears not only as the bearer of velocities, kinetic energy, and mechanical pressures, but also as the bearer of electromagnetic fields. Since such fields also occur in vacuo i.e. in free ether the ether also appears as bearer of electromagnetic fields. The ether appears indistinguishable in its functions from ordinary matter. Within matter it takes part in the motion of matter and in empty space it has everywhere a velocity; so that the ether has a definitely assigned velocity throughout the whole of space. There is no fundamental difference between Hertz's ether and ponderable matter (which in part subsists in the ether).

The Hertz theory suffered not only from the defect of ascribing to matter and ether, on the one hand mechanical states, and on the other hand electrical states, which do not stand in any conceivable relation to each other; it was also at variance with the result of Fizeau's important experiment on the velocity of the propagation of light in moving fluids, and with other established experimental results.

Such was the state of things when H. A. Lorentz entered upon the scene. He brought theory into harmony with experience by means of a wonderful simplification of theoretical principles. He achieved this, the most important advance in the theory of electricity since Maxwell, by taking from ether its mechanical, and from matter its electromagnetic qualities. As in empty space, so too in the interior of material bodies, the ether, and not matter viewed atomistically, was exclusively the seat of electromagnetic fields. According to Lorentz the elementary particles of matter alone are capable of carrying out movements; their electromagnetic activity is entirely confined to the carrying of electric charges. Thus Lorentz succeeded in reducing all electromagnetic happenings to Maxwell's equations for free space.

As to the mechanical nature of the Lorentzian ether, it may be said of it, in a somewhat playful spirit, that immobility is the only mechanical property of which it has not been deprived by H. A. Lorentz. 1t may be added that the whole change in the conception of the ether which the special theory of relativity brought about, consisted in taking away from the ether its last mechanical quality, namely, its immobility. How this is to be understood will forthwith be expounded.

The space-time theory and the kinematics of the special theory of relativity were modelled on the Maxwell-Lorentz theory of the electromagnetic field. This theory therefore satisfies the conditions of the special theory of relativity, but when viewed from the latter it acquires a novel aspect. For if K be a system of co-ordinates relatively to which the Lorentzian ether is at rest, the Maxwell-Lorentz equations are valid primarily with reference to K. But by the special theory of relativity the same equations without any change of meaning also hold in relation to any new system of co-ordinates K' which is moving in uniform translation relatively to K. Now comes the anxious question: Why must I in the theory distinguish the K system above all K' systems, which are physically equivalent to it in all respects, by assuming that the ether is at rest relatively to the K system? For the theoretician such an asymmetry in the theoretical structure, with no corresponding asymmetry in the system of experience, is intolerable. If we assume the ether to be at rest relatively to K, but in motion relatively to K', the physical equivalence of K and K' seems to me from the logical standpoint, not indeed downright incorrect, but nevertheless inacceptable.

The next position which it was possible to take up in face of this state of things appeared to be the following. The ether does not exist at all. The electromagnetic fields are not states of a medium, and are not bound down to any bearer, but they are independent realities which are not reducible to anything else, exactly like the atoms of ponderable matter. This conception suggests itself the more readily as, according to Lorentz's theory, electromagnetic radiation, like ponderable matter, brings impulse and energy with it, and as, according to the special theory of relativity, both matter and radiation are but special forms of distributed energy, ponderable mass losing its isolation and appearing as a special form of energy.

More careful reflection teaches us, however, that the special theory of relativity does not compel us to deny ether. We may assume the existence of an ether,; only we must give up ascribing a definite state of motion to it, i.e. we must by abstraction take from it the last mechanical characteristic which Lorentz had still left it. We shall see later that this point of view, the conceivability of which shall at once endeavour to make more intelligible by a somewhat halting comparison, is justified by the results of the general theory of relativity.

Think of waves on the surface of water. Here we can describe two entirely different things. Either we may observe how the undulatory surface forming the boundary between water and air alters in the course of time; or else with the help of small floats, for instance we can observe how the position of the separate particles of water alters in the course of time. If the existence of such floats for tracking the motion of the particles of a fluid were a fundamental impossibility in physics if, in fact, nothing else whatever were observable than the shape of the space occupied by the water as it varies in time, we should have no ground for the assumption that water consists of inovable particles. But all the same we could characterise it as a medium.

We have something like this in the electromagnetic field. For we may picture the field to ourselves as consisting of lines of force. If we wish to interpret these lines of force to ourselves as something inaterial in the ordinary sense, we are tempted to interpret the dynamic processes as motions of these lines of force, such that each separate line of force is tracked through the course of time. It is well known, however, that this way of regarding the electromagnetic field leads to contradictions.

Generalising we must say this: There inay be supposed to be extended physical objects to which the idea of motion cannot be applied. They may not be thought of as consisting of particles which allow themselves to be separately tracked through time. In Minkowski's idiom this is expressed as follows: Not every extended conformation in the four-dimensional world can be regarded as composed of worldthreads. The special theory of relativity forbids us to assume the ether to consist of particles observable through time, but the hypothesis of ether in itself is not in conflict with the special theory of relativity. Only we must be on our guard against ascribing a state of motion to the ether.

Certainly, from the standpoint of the special theory of relativity, the ether hypothesis appears at first to be an empty hypothesis. 1n the equations of the electromagnetic field there occur, in addition to the densities of the electric charge, only the intensities of the field. The career of electromagnetic processes in vacuo appears to be completely determined by tliese equations, uninfluenced by other physical quantities. The electromagnetic fields appear as ultimate, irreducible realities, and at first it seems superfluous to postulate a homogeneous, isotropic ether-medium, and to envisage electromagnetic fields as states of this medium.

But on the other hand there is a weighty argument to be adduced in favour of the ether hypothesis. To deny the ether is ultimately to assume that empty space has no physical qualities whatever. The fundamental facts of mechanics do not harmonize with this view. For the mechanical behaviour of a corporeal system hovering freely in empty space depends not only on relative positions (distances) and relative velocities, but also on its state of rotation, which physically may be taken as a characteristic not appertaining to the system in itself. In order to be able to look upon the rotation of the system, at least formally, as something real, Newton objectivises space. Since he classes his absolute space together with real things, for him rotation relative to an absolute space is also something real. Newton might no less well have called his absolute space ``Ether''; what is essential is merely that besides observable objects, another thing, which is not perceptible, inust be looked upon as real, to enable acceleration or rotation to be looked upon as something real.

It is true that Mach tried to avoid having to accept as real something which is not observable by endeavouring to substitute in inechanics a mean acceleration with reference to the totality of the

masses in the universe in place of an acceleration with reference to absolute space. But inertial resistance opposed to relative acceleration of distant masses presupposes action at a distance; and as the modern physicist does not believe that he may accept this action at a distance, he comes back once inore, if he follows Mach, to the ether, which has to serve as medium for the effects of inertia. But this conception of the ether to which we are led by Mach's way of thinking differs essentially from the ether as conceived by Newton, by Fresnel, and by Lorentz. Mach's ether not only conditions the behaviour of inert masses, but is also conditioned in its state by them.

Mach's idea finds its full development in the ether of the general theory of relativity. According to this theory the metrical qualities of the continuum of space-time differ in the environment of different points of space-time, and are partly conditioned by the matter existing outside of the territory under consideration. This space-time variability of the reciprocal relations of the standards of space and time, or, perhaps, the recognition of the fact that ``empty space'' in its physical relation is neither homogeneous nor isotropic, compelling us to describe its state by ten functions (the gravitation potentials g), has, I think, finally disposed of the view that space is physically empty. But therewith the conception of the ether has again acquired an intelligible content, although this content differs widely from that of the ether of the mechanical undulatory theory of light. The ether of the general theory of relativity is a medium which is itself devoid of all mechanical and kinematical qualities, but helps to determine mechanical (and electromagnetic) events.

What is fundamentally new in the ether of the general theory of relativity as opposed to the ether of Lorentz consists in this, that the state of the former is at every place determined by connections with the matter and the state of the ether in neighbouring places, which are amenable to law in the form of differential equations,; whereas the state of the Lorentzian ether in the absence of electromagnetic fields is conditioned by nothing outside itself, and is everywhere the same. The ether of the general theory of relativity is transmuted conceptually into the ether of Lorentz if we substitute constants for the functions of space which describe the former, disregarding the causes which condition its state. Thus we may also say, I think, that the ether of the general theory of relativity is the outcome of the Lorentzian ether, through relativation.

As to the part which the new ether is to play in the physics of the future we are not yet clear. We know that it determines the metrical relations in the space-time continuum, e.g. the configurative possibilities of solid bodies as well as the gravitational fields; but we do not know whether it has an essential share in the structure of the electrical elementary particles constituting matter. Nor do we know whether it is only in the proximity of ponderable masses that its structure differs essentially from that of the Lorentzian ether; whether the geometry of spaces of cosmic extent is approximately Euclidean. But we can assert by reason of the relativistic equations of gravitation that there must be a departure from Euclidean relations, with spaces of cosmic order of magnitude, if there exists a positive mean density, no matter how small, of the matter in the universe. In this case the universe must of necessity be spatially unbounded and of finite magnitude, its inagnitude being determined by the value of that inean density.

If we consider the gravitational field and the electromagnetic field from the standpoint of the ether hypothesis, we find a remarkable difference between the two. There can be no space nor any part of space without gravitational potentials; for these confer upon space its metrical qualities, without which it cannot be imagined at all. The existence of the gravitational field is inseparably bound up with the existence of space. On the other hand a part of space may very

well be imagined without an electromagnetic field; thus in contrast with the gravitational field, the electromagnetic field seems to be only secondarily linked to the ether, the formal nature of the electromagnetic field being as yet in no way determined by that of gravitational ether. From the present state of theory it looks as if the electromagnetic field, as opposed to the gravitational field, rests upon an entirely new formal motif, as though nature might just as well have endowed the gravitational ether with fields of quite another type, for example, with fields of a scalar potential, instead of fields of the electromagnetic type.

Since according to our present conceptions the elementary particles of matter are also, in their essence, nothing else than condensations of the electromagnctic field, our present view of the universe presents two realities which are completely separated from each other conceptually, although connected causally, namely, gravitational ether and electromagnetic field, or as they might also be called space and matter.

Of course it would be a great advance if we could succeed in comprehending the gravitational field and the electromagnetic field together as one unified conformation. Then for the first time the epoch of theoretical physics founded by Faraday and Maxwell would reach a satisfactory conclusion. The contrast between ether and matter would fade away, and, through the general theory of relativity, the whole of physics would become a complete system of thought, like geometry, kinematics, and the theory of gravitation. An exceedingly ingenious attempt in this direction has been made by the mathematician H. Weyl,; but I do not believe that his theory will hold its ground in relation to reality. Further, in contemplating the immediate future of theoretical physics we ought not unconditionally to reject the possibility that the facts comprised in the quantum theory may set bounds to the field theory beyond which it cannot pass.

Recapitulating, we may say that according to the general theory of relativity space is endowed with physical qualities; in this sense, therefore, there exists an ether. According to the general theory of relativity space without ether is unthinkable; for in such space there not only wonld be no propagation of light, but also no possibility of existence for standards of space and time (measuring-rods and clocks), nor therefore any space-time intervals in the physical sense. But this ether may not be thought of as endowed with the quality characteristic of ponderable inedia, as consisting of parts which may be tracked through time. The idea of motion may not be applied to it.