

The Outstanding Problems of Relativity.<sup>1</sup>

By Prof. E. T. WHITTAKER, F.R.S.

IT was in January 1914 that Einstein (*Zeits. f. Math. u. Phys.*, 63, p. 215; 1914) made his great departure from the Newtonian doctrine of gravitation by abandoning the idea that the gravitational potential is scalar. The thirteen eventful years which have passed since then have seen the rapid development of the new theory, which is called General Relativity, and the confirmation by astronomers and astrophysicists of its predictions regarding the bending of light-rays by the sun and the displacement of spectral lines. At the same time a number of new problems have arisen in connexion with it; and perhaps the time has now come to review the whole situation and to indicate where there is need for further investigation.

Prof. G. F. FitzGerald of Dublin long held an opinion which he expressed in 1894 in the words: "Gravity is probably due to a change of structure of the ether, produced by the presence of matter" (FitzGerald's "Scientific Writings," p. 313). Perhaps this is the best description of Einstein's theory that can be given in a single sentence in the language of the older physics: at any rate it indicates the three salient principles, first, that gravity is not a force acting at a distance, but an effect due to the modification of space (or, as FitzGerald would say, of the ether) in the immediate neighbourhood of the body acted on; secondly, that this modification is propagated from point to point of space, being ultimately connected in a definite way with the presence of material bodies; and thirdly, that the modification is not necessarily of a scalar character. The mention of the ether would be criticised by many people to-day as something out-of-date and explicable only by the circumstance that FitzGerald was writing thirty-three years ago; but even this criticism will not be universal; for Wicchert and his followers have actually combined the old ether theory with ideas resembling Einstein's by the hypothesis that gravitational potential is an expression of what we may call the specific inductive capacity and permeability of the ether, these qualities being affected by the presence of gravitating bodies. Assuming that matter is electrical in its nature, it is inferred that matter will be attracted to places of greater dielectric constant. It seems possible that something of this sort was what FitzGerald had in mind.

Let us now consider some of the consequences of Einstein's theory. One of the first of them is that when a planet moves round a central attracting body in a nearly circular orbit, the perihelion of the orbit advances by (approximately)  $6\pi v^2/c^2$  in each revolution, where  $v$  is the planet's velocity and  $c$  is the velocity of light. This gives for the motion of the perihelion of Mercury almost exactly the amount (42" per century) which is found from observation. Another consequence is that light-rays which pass near a massive body are deflected, the bending at the sun's limb being 1".75. This

was confirmed observationally by the British expeditions to the eclipse of May 1919, and still more decisively by the Lick Observatory expedition to the Australian eclipse of September 1922: the Lick observers found for the shift  $1".72 \pm 0".11$ , which differs from Einstein's predicted value by much less than its estimated probable error. Yet another result of general relativity is that, by the Principle of Equivalence, light which reaches us from a place of different gravitational potential (such as the sun) must exhibit a kind of Doppler effect. This 'gravitational shift of the solar spectral lines' is now generally admitted to be confirmed by comparisons of wave-lengths at the centre of the sun's disc with wave-lengths from the arc *in vacuo*; and in 1925 the effect was observed, on a much larger scale, by W. S. Adams in the spectrum of the companion of Sirius.

Besides the effects which have been verified observationally, there are many consequences of Einstein's theory which are of interest as opening up new fields or presenting new inter-relations of phenomena in astronomy and physics. For example, there is a contribution to the precession of the equinoxes which, unlike ordinary precession, does not depend on the oblateness of the earth. Again, the bending of the rays of light near a gravitating body, which has been observed in the case of the sun and the companion of Sirius, may, theoretically at any rate, be so pronounced that the ray is permanently captured by the attracting body, and describes for ever a track round and round it, which approaches spirally and asymptotically to a circle the centre of which is at the centre of gravitation. Yet another deduction is that an electrified body, or a single electron, which is at rest in a varying gravitational field, must emit radiation. Indeed, now that a definite connexion has been set up between electricity and gravitation, the whole of electromagnetic theory must be rewritten.

As a further illustration of the (as yet) unexplored possibilities of the new physics, let us consider the well-known equations for the potential of Newtonian gravitation, namely, Laplace's equation in space where there is no matter, and Poisson's equation in space where matter of density  $\rho$  is present. In general relativity, when the field is statical, these are replaced by an equation which reduces to Laplace's equation in one extreme case (when no matter or energy is present at the point) and to Poisson's equation in another extreme case (when the energy is entirely in the form of ordinary matter), but it offers an infinite variety of possibilities intermediate between the two, in which energy is present but not in the form of ordinary matter. It is possible that this equation, which evidently suggests an approach to the new wave-mechanics, may play as important a part in the microphysics and astrophysics of the future as the equations of Laplace and Poisson have played in the ordinary physics of the past.

<sup>1</sup> From the presidential address to Section A (Mathematical and Physical Sciences) of the British Association delivered at Leeds on Sept. 5.

The fundamental researches of Einstein and Hilbert, with the discovery of the field equations of gravitation, were published in 1915. At that time German scientific journals did not reach Great Britain regularly, and British physicists and mathematicians were mostly occupied in one way or another with duties arising out of the Great War; so that comparatively little notice was taken of the new theory on this side of the North Sea during the first year or two of its existence, and indeed it was not until the end of the War that most of us had any opportunity of studying it. In Germany, however, it was quickly realised that general relativity was one of the most profound and far-reaching contributions that had ever been made to science. Its successful prediction of new phenomena of a most unexpected kind was an event of the first importance, but still more significant was its complete subversion of the foundations of physics and reconstruction of the whole subject on a new basis. From time immemorial the physicist and the pure mathematician had worked on a certain agreement as to the shares which they were respectively to take in the study of Nature. The mathematician was to come first and analyse the properties of space and time, building up the primary science of geometry; then, when the stage had thus been prepared, the physicist was to come along with the *dramatis personæ*—material bodies, magnets, electric charges, light, and so forth—and the play was to begin. But in Einstein's revolutionary conception, the characters created the stage as they walked about on it: geometry was no longer antecedent to physics, but indissolubly fused with it into a single discipline. The properties of space, in general relativity, depend on the material bodies that are present; Euclidean geometry is deposed from its old position of priority, and from acceptance as a valid representation of space; indeed its whole spirit is declared to be alien to that of modern physics, for it attempts to set up relations between points which are at a finite distance apart, and thus is essentially an action-at-a-distance theory; and in the new world no direct relations exist at all except between elements that are contiguous to each other.

The scheme of general relativity, as put forward by Einstein in 1915, met with some criticism as regards the unsatisfactory position occupied in it by electrical phenomena. While gravitation was completely fused with metric, so that the notion of a mechanical force on ponderable bodies due to gravitation attraction was completely abolished, the notion of a mechanical force acting on electrified or magnetised bodies placed in an electric or magnetic field still persisted as in the old physics. This seemed to be an imperfection, and it was felt that sooner or later everything, including electromagnetism, would be re-interpreted and represented in some way as consequences of the pure geometry of space and time. In 1918, Weyl proposed to effect this by rebuilding geometry once more on a new foundation. He devised a geometry more general than the Riemannian geometry which had

been adopted by Einstein; instead of being specified, like the Riemannian geometry, by a single quadratic differential form, it is specified by a quadratic differential form and a linear differential form together. Thus he succeeded in exhibiting both gravitation and electricity as effects of the metric of the world.

The enlargement of geometrical ideas thus achieved was soon followed by still wider extensions of the same character, due to Eddington, Schouten, Wirtinger, and others. From the point of view of the geometer, they constituted striking and valuable advances in his subject, and they seemed to offer an attractive prospect to the physicist of combining the whole of our knowledge of the material universe into a single unified theory. The working out of the various possible alternative schemes for identifying these more general geometries with physics has been the chief occupation of relativists during the last nine years. Many ingenious proposals and adaptations have been published, and more than one author has triumphantly announced that at last the problem has been solved. But I do not think that any of the theories can be regarded as satisfactory, and within the last year or two a note of doubt has been perceptible; Were we after all on the right track? At last Einstein himself (*Math. Ann.*, 97, p. 99; 1926) has made up his mind and renounced the whole movement. The present position, then, is that the years 1918–1926 have been spent chiefly in researches which, while they have contributed greatly to the progress of geometry, have been on altogether wrong lines so far as physics is concerned; and we have now to go back to the pre-1918 position and make a fresh start, with the definite conviction that the geometry of space-time is Riemannian.

Granting, then, this fundamental understanding, we have now to inquire into the axiomatics of the theory. This part of the subject has received less attention in Great Britain than elsewhere, perhaps because of the more or less accidental circumstance that the most prominent and distinguished exponents of relativity in England happened to be men whose work lay in the field of physics and astronomy rather than in mathematics, and who were not specially interested in questions of logic and rigour. It is, however, evidently of the highest importance that we should know exactly what assumptions must be made in order to deduce our equations, especially since the subject is still in a rather fluid condition.

What we want to do, then, is to set forth the axiomatics of general relativity in the same form as we have been accustomed to give to the axiomatics of any other kind of geometry—that is, to enunciate the primitive or undefined concepts, then the definitions, the axioms, and the existence-theorems, and lastly the deductions. In the course of the work we must prove that the axioms are compatible with each other, and that no one of them is superfluous.

The usual way of introducing relativity is to talk about measuring-rods and clocks. This is, I think,

a very natural and proper way of introducing the doctrine known as 'special relativity,' which grew out of FitzGerald's hypothesis of the contraction of moving bodies, and was first clearly stated by Poincaré in 1904, and further developed by Einstein in 1905. But general relativity, which came ten years later, is a very different theory. In general relativity there are no such things as rigid bodies—that is, bodies for which the mutual distance of every pair of particles remains unaltered when the body moves in the gravitational field. That being so, it seems desirable to avoid everything akin to a rigid body—such, for example, as measuring-rods or clocks—when we are laying down the axioms of the subject. The axioms should obviously deal only with the simplest constituents of the universe. Now if one of my clocks or watches goes wrong, I do not venture to try to mend it myself, but take it to a professional clockmaker, and even he is not always wholly successful, which seems to me to indicate that a clock is not one of the simplest constituents of the universe. Some of the expounders of relativity have recognised the existence of this difficulty, and have tried to turn it by giving up the ordinary material clock with its elaborate mechanism, and putting forward in its place what they call an atomic clock; by which they mean a single atom in a gas, emitting light of definite frequency. Unfortunately, the atom is apparently quite as complicated in its working as a material clock, perhaps more so, and is less understood; and the statement that the frequency is the same under all conditions, whatever is happening to the atom, is (whether true or not) a highly complex assumption which could scarcely be used in an axiomatic treatment of the subject until it has been dissected into a considerable number of elementary axioms, some of them perhaps of a disputable character.

It seems to me that we should abandon measuring-rods and accurate clocks altogether, and begin with something more primitive. Let us then take any system of reference for events—a network of points to each of which three numbers are assigned—which can serve as spatial co-ordinates, and a number indicating the succession of events at each point to serve as a temporal co-ordinate. Let us now refer to this co-ordinate system the paths which are traced by infinitesimal particles moving freely in the gravitational field. Then it is one of the fundamental assumptions of the theory that these paths are the geodesics belonging to a certain quadratic differential form. The truth or falsity of this assumption may, in theory at any rate, be tested by observation, since if the paths are geodesics they must satisfy certain purely geometrical conditions, and whether they do or not is a question to be settled by experience.

Granting for the present that the paths do satisfy these conditions, let us inquire if a knowledge of the paths or geodesics is sufficient to enable us to determine the quadratic form. The answer to this is in the negative, as may easily be seen if we consider for a moment the non-Euclidean geometry defined by a Cayley-Klein metric in

three-dimensional space. In the Cayley-Klein geometry the geodesics are the straight lines of the space; but a knowledge of this fact is not sufficient to determine the metric, since the Absolute may be any arbitrary quadric surface.

In order to determine the quadratic form in general relativity we must, then, be furnished with some information besides the knowledge of the paths of material particles. It is sufficient, as Levi-Civita has remarked, that we should be given the null geodesics, *i.e.* the geodesics along which the quadratic form vanishes. In the Cayley-Klein geometry these are the tangents to the Absolute; in general relativity they are tracks of rays of light.

So from our knowledge of the paths of material particles and the tracks of rays of light we can construct the particular quadratic form, and then we are ready for the next great axiom, namely, Einstein's Principle of Covariance. The point I wish specially to make is that in the above treatment there has been no mention either of length or of time: neither measuring-rod nor clock has been introduced in any way. We have left open the question whether the quadratic form does or does not represent anything which can be given directly by measuring-rods and clocks. For my own part I incline to think that the notions of length of material bodies, and time of clocks, are really rather complex notions which do not normally occur in the early chapters of axiomatic physics. The results of the ether-drift experiments of D. C. Miller at Mount Wilson in 1925, if confirmed, would seem to indicate that the geometry which is based on rigid measuring-rods is actually different from the geometry which is based on geodesics and light-rays.

The actual laws of Nature are most naturally derived, it seems to me, from the Minimum Principle enunciated in 1915 by Hilbert. This principle is the grand culmination of the movement begun 2000 years ago by Hero of Alexandria with his discovery that reflected light meets the mirror at a point such that the total path between the source of light and the eye is the shortest possible. In the seventeenth century, Hero's theorem was generalised by Fermat into his "Principle of Least Time" that "Nature always acts by the shortest course," which suffices for the solution of all problems in geometrical optics. A hundred years later, this was further extended by Maupertuis, Euler, and Lagrange, into a general principle of 'Least Action' of dynamical systems, and in 1834, Hamilton formulated his famous Principle which was capable of reducing all the known laws of Nature—gravitational, dynamical, and electrical—to a representation as minimum-problems.

Hilbert's minimum principle in general relativity is a direct application of Hamilton's principle, in which the contribution made by gravitation is the integral of the Riemann scalar curvature. Thus gravitation acts so as to make the total amount of the curvature of space-time a minimum: or, as we may say, *gravitation simply represents a continual effort of the universe to straighten itself out.* This is general relativity in a single sentence.

I have already explained that the curvature of space-time at any point at any instant depends on the physical events that are taking place there: in statical systems, where we can consider space of three dimensions separately from time, the mean curvature (*i.e.* the sum of the three principal curvatures) of the space at any point is proportional to the energy-density at the point. Since, then, the curvature of space is wholly governed by physical phenomena, the suggestion presents itself that the metric of space-time may be determined *wholly* by the masses and energy present in the universe, so that space-time cannot exist at all except in so far as it is due to the existence of matter. This doctrine, which is substantially due to Mach, was adopted in 1917 by Einstein, and has led to some interesting developments. The point of issue may be illustrated by the following concrete problem: If all matter were annihilated except one particle which is to be used as a test-body, would this particle have inertia or not? The view of Mach and Einstein is that it would not; and in support of this view it may be urged that, according to the deductions of general relativity, the inertia of a body is increased when it is in the neighbourhood of other large masses; it seems needless, therefore, to postulate other sources of inertia, and simplest to suppose that *all* inertia is due to the presence of other masses. When we confront this hypothesis with the facts of observation, however, it seems clear that the masses we know to exist—the solar systems, stars, and nebulae—are insufficient to confer on terrestrial bodies the inertia which they actually possess; and, therefore, if Mach's principle were adopted, it would be necessary to postulate the existence of enormous quantities of matter in the universe which have not been detected by astronomical observation, and are called into being simply in order to account for inertia in other bodies. This is, after all, no better than regarding some part of inertia as intrinsic.

Under the influence of Mach's doctrine, Einstein made an important modification of the field-equations of gravitation. He now objected to his original equations of 1915 on the ground that they possessed a solution even when the universe was supposed void of matter, and he added a term—the 'cosmological term' as it is called—with the idea of making such a solution impossible. After a time it was found that the new term did not do what it had been intended to do, for the modified field-equations still possessed a solution—the celebrated 'De Sitter world'—even when no matter was present; but the 'De Sitter world' was found to be so excellent an addition to the theory that it was adopted permanently, and with it of course the 'cosmological term' in the field-equations; so that this term has been retained for exactly the opposite reason to that for which it was originally introduced.

The 'De Sitter world' is simply the universe as it would be if all minor irregularities were smoothed out: just as when we say that the earth is a spheroid, we mean that the earth would be a

spheroid if all mountains were levelled and valleys filled up. In the case of the De Sitter universe, the levelling is a more formidable operation, since we have to smooth out the earth, the sun, and all the heavenly bodies, and reduce the world to a complete uniformity. But after all, only a very small fraction of the cosmos is occupied by material bodies; and it is interesting to inquire what space-time as a whole is like when we simply ignore them.

The answer is, as we should expect, that it is a manifold of constant curvature. This means that it is isotropic (*i.e.* the Riemann curvature is the same for all orientations at the same point), and is also homogeneous. As a matter of fact, there is a well-known theorem that any manifold which is isotropic in this sense is necessarily also homogeneous, so that the two properties are connected. A manifold of constant curvature is a projective manifold, *i.e.* ordinary projective geometry is valid in it when we regard geodesics as straight lines; and it is possible to move about in it any system of points, discrete or continuous, rigidly, *i.e.* so that the mutual distances are unaltered.

We are thus brought to the question of the dimensions of the universe: What is the length of the complete straight line, the circuit of all space? The answer must be furnished by astrophysical observations, interpreted by a proposition which belongs to the theory of 'De Sitter's world,' namely, that the lines of the spectrum of a very distant star should be systematically displaced; the amount of displacement is proportional to the ratio of the distance of the star from the observer to the constant radius of curvature  $R$  of the universe. In attempting to obtain the value of  $R$  from this formula we meet with many difficulties: the effect is entangled with the ordinary Doppler effect due to the radial velocity of the star; it could in any case only be of appreciable magnitude with the most distant objects, and there is the most serious difference of opinion among astronomers as to what the distance of these objects really is. Within the last twelve months the distance of the spiral nebula M 33 Trianguli has been estimated by Dr. Hubble, of the Mount Wilson Observatory, at 857,000 light-years, and by Dr. Perrine, the Director of the Cordoba Observatory, at only 30,000 light-years; and there is a similar uncertainty of many thousands per cent. in regard to all other very remote objects.

In these circumstances we hesitate to assign a definite length for the radius of curvature of the universe; but it is millions of light-years, though probably not greater than about a hundred millions. The curvature of space at any particular place due to the general curvature of the universe is therefore quite small compared to the curvature which may be imposed on it locally by the presence of energy. By a strong magnetic field we can produce a curvature with a radius of only 100 light-years, and of course in the presence of matter the curvature is far stronger still. So the universe is like the earth, on which the local curvature of hills and valleys is far greater than the general curvature of the terrestrial globe.