425

THE

HIBBERT JOURNAL

EUCLID, NEWTON, AND EINSTEIN.

C. D. BROAD, M.A.

I PROPOSE to try in the present paper to put into simple terms, which shall neither make a layman feel dizzy nor a mathematician feel sick, the main points of Einstein's principle of relativity. The bright young men who instruct us daily on the problems of science and religion from their desks in newspaper offices have succeeded in conveying the impression that Einstein has proved Euclid to be a fool and Newton an ignoramus. One famous journal, with a large circle of readers in Scotland, so far departed from the customary subject of its leading articles—the reiterated demand that the ex-Kaiser shall be executed, condemned, and tried—as to inform the world, on the alleged authority of one who obstinately insists that he is a German in spite of the most charitable attempts to prove him to be Swiss, that "circles are not round." It is rash, but not (as yet) criminal, to suggest that, when newspaper editors write on subjects which they cannot be expected to understand, they are liable to talk nonsense. I therefore venture to think that there may still be something useful to be said as to the precise relation of Einstein's theories to Euclidean geometry and Newtonian physics. It will also be worth while to inquire whether the new views have any important bearing on philosophical problems. It is more than likely that I shall fail to fulfil my promises; but I can hardly be more absurd than the newspapers or less intelligible than the experts.

It is highly important at the outset to understand clearly that Einstein has put forward two theories of relativity: the "restricted theory," as he calls it, and the present generalised theory. The former has nothing whatever to do with gravitation; the latter has no *specially* intimate connection with optical or electrical phenomena. The new theory is, in a very important sense, an extension of the old one; but it is not a mere generalisation which contains the old one unmodified as a part. *E.g.* the constancy of the velocity of light is the keystone of the old theory, whilst this velocity is not absolutely constant on the new one. I shall begin with a sketch of the restricted principle, which has been before the world since 1906.

The grounds for the restricted theory are best understood by considering the famous Michelsen-Morley experiment. The principles of this are perfectly easy to grasp. Suppose one had a platform moving through the ether in a certain direction with a constant velocity. On this platform let there be an observer, a source of light, and a couple of mirrors. Draw a line on the platform through the source of light and parallel to the direction in which it is moving. Draw another line on the platform through the source at right angles to the first. Mark off equal distances from the source on both these lines. At the points thus obtained place the two mirrors normally to each line respectively. At a certain moment let the source give out a flash of light. The part of this that travels parallel to the direction in which the platform is moving will have to travel more than the marked distance before it reaches the mirror; for the mirror will have moved on through the ether while the light is travelling up to it, and thus the light will be overtaking it. Now let it be reflected back along its old path. It will now have to travel less than the marked distance, because, while it travels back through the ether, the source will be moving up to meet it. The total distance travelled by the light through the ether from the time when it leaves the source to the time when it gets back to it can easily be shown to be $\frac{2lc^2}{c^2-v^2}$, where l is the marked length, c the velocity of light relative to the ether, and v the velocity of the platform relative to the ether.

Let us now follow the fate of the light that travels to the other mirror and is reflected back from it to the source. By the time that such light gets back the source is no longer at its original position in the ether. Hence light that travelled out at right angles to the direction in which the platform is moving would not return to the source at all after reflexion, for it would be returned, not to the place where the source *is*, but to where it *was* when the flash left it. We have therefore to consider light which strikes the mirror at a point in the ether equidistant from the point at which the source was when the light left it and the point at which it will be when the light returns to it. The light thus describes an isosceles triangle in the ether, with a point on the mirror as apex and the distance between the initial and final positions of the source as base. It can be shown that the actual distance travelled in the ether by such light between leaving the source and returning to it is $\frac{2lc}{\sqrt{c^2-v^2}}$.

In fact, the light that travelled to the first mirror and back has traversed a distance $\frac{2l}{1-\frac{c^2}{v^2}}$, and the light that travelled to the second and back has traversed a distance $\frac{2l}{\sqrt{1-\frac{v^2}{c^2}}}$. Now, when beams of light from a common source come to the same point after travelling different distances, they "interfere," *i.e.* they produce a spectrum with bright and dark bands. And there is a perfectly definite relation between the position of these bands and the difference in the distance travelled by the two beams. Hence in the present case there should be interference, and it should be possible to determine from it the velocity of the platform with respect to the ether, if our instruments be delicate enough. In the Michelsen-Morley experiment the platform was the earth and v was the velocity of the earth in its orbit. The apparatus was quite delicate enough to detect effects of the order of magnitude predicted. *No trace of such effects was found.* There are many other experiments, more difficult to understand without special knowledge of the laws of electromagnetism, by which one might hope to detect and measure the velocity of the earth with respect to the ether. *In no case has any such effect been observed*, though the methods used were quite delicate enough to detect them if they had been present. This negative experimental fact, that no effect due to the uniform motion of a body through the ether has been observed, although it was predicted and although it could have been noticed and measured if present, is the basis of the restricted theory of relativity.

These being the facts, what conclusions are we to base on them? In the first place, what assumptions did we make when we calculated the different distances travelled by the two beams? Apart from assumptions about the measurement of space and time, with which we shall have to deal later, we assumed (*a*) that the ether is not dragged along in any way by the platform, as water would be by a stick moved through it; (*b*) that the velocity of light in stagnant ether is the same in all directions; and (*c*) that the fact that a source which emits light is itself in motion does not affect the velocity of the emitted light. Would it be reasonable to account for the negative result of the Michelsen-Morley experiment by rejecting or modifying any of these assumptions? As regards (*a*), the opposite alternative would bring us in conflict with another set of experimental facts. The aberration of light from a star due to the annual motion of the earth will be different according to whether the ether is stagnant or whether the earth drags some of it along in its course. The amount of aberration that should be observed on any hypothesis of this nature can be calculated and compared with that which is actually found. What is actually found is that which would follow from the hypothesis of a stagnant ether. If we assume that the earth drags the ether along with it to the extent needed to account for the negative result of the Michelsen-Morley experiment, the resulting value of the aberration will differ widely from that which is actually found. Hence assumption (*a*) cannot be rejected.

The assumption (*b*) seems to be the only reasonable one to make on the subject. Nor would it help us to reject it, for, since the earth is moving in its orbit in different directions in the ether at different times of the year, the supposition that the velocity of light in the ether varies with absolute direction in the ether, even if it can be made intelligible, ought at most to make the result of the Michelsen-Morley experiment null at one point in the earth's orbit. It should make the discrepancy between prediction and observation worse than before at other seasons of the year.

(*c*) On the wave-theory of light there is no reason why the velocity of a source at the moment of emission should have any effect on the velocity with which the disturbance set up subsequently travels in the ether. If we held the corpuscular theory of light, matters would be different; for a corpuscle shot out of a body that was itself moving would presumably have a velocity compounded of that due to the discharging impulse and that of the source itself. But the evidence for the wave-theory and against the corpuscular theory is so strong that it seems idle to try to explain the negative result of the experiment by making an assumption which is only plausible on the latter hypothesis.

It seems, then that all the *physical* assumptions that led us to expect a positive result from the Michelsen-Morley experiment are highly plausible, and that the rejection of any of them will merely bring us into conflict with some other set of well-attested experimental facts. We are thus absolutely forced to turn our attention to the assumptions that have been made as to the measurement of distances and time-lapses. This brings us, as regards space, to the celebrated Lorentz-Fitzgerald Contraction, and, as regards time, to the notion of Local Time.

It will be remembered that we marked out two lines on our platform, both passing through the source, one parallel to the direction in which the platform is moving and the other at right angles to this. Along these we measured off what we took to be the same distance l. On the assumption that we had really measured the same distance along both lines, we saw that the distance travelled by the light which goes parallel to the motion of the platform is $\dfrac{2l}{1-\dfrac{v^2}{c^2}}$, whilst that travelled by the light which goes at right angles to this direction is $\dfrac{2l}{\sqrt{1-\dfrac{v^2}{c^2}}}$.

Yet nothing corresponding to this difference can be observed. We could account for this fact if the distance at right angles to the direction of motion which is measured as l "really is" l, whilst the distance parallel to the direction of motion which is measured as l "really" is only $l\sqrt{1-\dfrac{v^2}{c^2}}$. The actual distance travelled by both beams will then be the same, viz. $\dfrac{2l}{\sqrt{1-\dfrac{v^2}{c^2}}}$, and therefore the negative result of the experiment will be explained. If we suppose that *everything* contracts in this proportion in the direction of its velocity with respect to the ether, it is obvious that no process of direct measurement will tell us of the fact. A measuring rod that goes l times into a line at right angles to the direction of motion will also go l times into a line of length $l\sqrt{1-\dfrac{v^2}{c^2}}$ parallel to the direction of motion, because its own length when turned in the latter direction will change from 1 to $\sqrt{1-\dfrac{v^2}{c^2}}$. This is what is called the Lorentz-Fitzgerald Contraction.

We can now deal with the question of Local Time. Let us suppose that the observer on our moving platform is trying to determine the velocity of light relative to the platform. The numerical value of a velocity will naturally depend on the units of space and time chosen; it will be different according to whether we reckon in centimetres and seconds or in inches and hours. Therefore, if we wish to compare the velocity of light relative to a moving platform with that in the stagnant ether we must be sure that our time-measurer is going at the same rate. We have supposed the velocity of light in stagnant ether to have a certain numerical value c when distances are measured in centimetres and time-lapses in seconds. Now, suppose we have some time-measurer at our source, and that the arrangements are otherwise as before. We first want to be sure that it is measuring seconds in order to get a fair comparison between the velocity of light relative to the platform and that in the stagnant ether. Now, we have seen that, assuming the Lorentz-Fitzgerald contraction, the distance *in the ether* travelled by a beam which leaves the source, strikes one of the mirrors, and then returns to the source, is $\dfrac{2l}{\sqrt{1-\dfrac{v^2}{c^2}}}$, where l is the measured distance from source to mirror in centimetres. Since c is assumed as the numerical value of the velocity of light in the ether, it is clear that our clock ought to indicate a lapse of $\dfrac{2l}{\sqrt{1-\dfrac{v^2}{c^2}}} \div c$ between the departure and return of the light, if it is accurately measuring time in seconds. For a fair comparison we should have to set it so as to do this. Now, the distance travelled by this beam *relatively to the platform* when measured in centimetres is $2l$. Therefore the velocity of light relatively to the platform will be $\dfrac{2l}{\dfrac{2l}{c\sqrt{1-\dfrac{v^2}{c^2}}}}$, or $c\sqrt{1-\dfrac{v^2}{c^2}}$ centimetres per second. It will thus vary with the velocity of the platform. This seems a perfectly reasonable result, and exactly what one might expect. But it is not confirmed by experience. Actually we find that the measured velocity of light does *not* depend on the velocity of the source, the observer, and his instruments. So we have another conflict between prediction and observation to account for. Evidently we cannot meet it by any further modifications about the measurement of *space*, or we shall have the Michelsen-Morley difficulty, which we had hoped to be safely buried, back on our hands. We are therefore forced to reconsider our measurements of *time*. Suppose that when a period of one second has "really elapsed" our clock indicates a lapse of $\sqrt{1-\frac{v^2}{c^2}}$ seconds, *i.e.* is a little slow. Then the *measured* lapse between the departure and the arrival of the beam at the source will not be $\frac{2l}{c\sqrt{1-\frac{v^2}{c^2}}}$ but $\frac{2l}{c}$ seconds. The measured distance traversed by the beam with respect to the platform is of course still $2l$ centimetres. Thus the measured velocity of light with respect to the platform now becomes $\frac{2l}{\frac{2l}{c}}$, *i.e.* c centimetres per second. It is thus independent of the velocity of the platform, which, as we saw, is the result actually observed. We have therefore to assume that the clock at the source goes slower than the same clock at the same place would do if the platform were at rest in the ether, and that the ratio is $\sqrt{1-\frac{v^2}{c^2}}:1$. This assumption, of course, makes no difference to the Lorentz-Fitzgerald solution of the Michelsen-Morley difficulty.

But we are not even yet at the end of our troubles about the measurement of time. We have assumed so far that we have only had to deal with *one* time-measurer in *one* place; for the light came back in the end to the place from which it started, and the lapse was measured by the clock there. This, of course, does accord with the way in which the velocity of light actually has been measured by purely terrestrial experiments such as those of Fizeau and Foucault. Still, it is clear that we often want to compare the time at which something leaves one place with that at which it arrives in *another* place, and that in order to be able to do this we must have some ground for believing that the clocks in the two places are going at the same rate and that they agree in their zeros. Now, the mere fact that they agreed in these respects when together does not guarantee that they will always continue to do so when one has been taken away to a distance. In the case of a pair of clocks, *e.g.*, the shaking that one of them gets on the journey, the possibly different average temperature of the region to which it is moved, the difference in the gravitational attraction at different parts of the earth, and many other factors, would make it most unsafe to argue that, because they agreed when together, they must continue to agree when widely separated. It is thus absolutely necessary to have some criterion of sameness of rate and sameness of zero which can be applied even when two clocks are at a great distance from each other. Now, the only criterion that suggests itself makes use of signals sent from the place where one is to the place where the other is. Let a light signal be sent from clock A when it marks t_A and received at clock B when this marks t_B. Let this be repeated when the first clock marks t_A', and let it be received at B when the second marks t_B'. Then if we find that $t_B' - t_B$, the lapse recorded between the reception of the two signals, is equal to $t_A' - t_A$, the lapse recorded between the despatch of them, it seems reasonable to conclude that the two clocks are going at the same rate. Again, if a signal leaves A at t_A, reaches B when the clock there marks t_B, is immediately reflected back to A, and reaches there when the clock at A marks t_A', it seems reasonable to conclude that the two clocks agree in their zeros provided that $t_B = \frac{1}{2}(t_A + t_A')$. The plain fact is that these criteria seem reasonable, that no others suggest themselves, and that *some* criterion is necessary if we are to deal at all with events that happen at different places. Moreover, with this criterion, and with this alone, will observers find the same value for the velocity of light, whether they measure it by observations all carried out at one place (as described above) or by observations made at two distant places. We can easily see this as follows. We have seen that the velocity of light relative to a platform, as determined by observations at a single place on the platform, will be found to be c, no matter what velocity the platform may have in the ether. Now, let our other clock be put where the first mirror is in the Michelsen-Morley experiment, *i.e.* at a measured distance l from the source in the direction of motion of the platform. Let a flash leave the source A when the clock there marks 0, reach the clock at B when this marks t_B, be reflected back, and return to A when the clock at A marks t_A. Then by our criterion $t_B = \frac{1}{2}(0 + t_A) = \frac{1}{2}t_A$. But we know that the velocity of light relative to the platform, as measured entirely by observations at A with the clock at A, is c. And the measured distance that this light has travelled relatively to the platform is $2l$, *i.e.* the distance on the platform backwards and forwards between A and B. Hence $t_A = \frac{2l}{c}$. Hence

t_B (which $= \frac{1}{2}t_A$) is $\frac{l}{c}$. That is to say, a beam of light, which left A when A's clock marked 0 and travelled the distance l relative to the platform to the point B, reaches B when the clock there marks a time $\frac{l}{c}$. Thus the observers at A and B on comparing notes will again conclude that the velocity of light with respect to the platform is c, which is exactly the same conclusion as experimenters who confined themselves to making observations at A with A's clock had already reached. So that the conventions for standardising distant clocks are not only reasonable in themselves, but are the only ones that will lead to the same measure of the velocity of light with respect to the platform when two different but equally reasonable methods of measuring that velocity are used.

But, as I shall now show, these conventions, reasonable as they are, imply that, if the platform is moving, clocks at different positions are wrong in comparison with what they would record if the platform were at rest, not merely in the sense already noticed that they are going at a slower rate, but also in the further sense that agreement between their readings does not imply identity of time. We have just seen that if a flash leaves A when the clock there reads 0, it will reach B when the clock *there* reads $\frac{l}{c}$. Now, if there were nothing wrong with the clock at B except the systematic slowness of rate that we have already noticed, the real lapse of time corresponding to a reading $\frac{l}{c}$ should be $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{l}{c}$. But actually the light that left A at 0 has travelled (i.) a distance $l\sqrt{1-\frac{v^2}{c^2}}$ in the ether (taking account of the Lorentz-Fitzgerald contraction), and (ii.) has had further to catch up B, which is itself moving in the same direction with velocity v. A very simple calculation will show that the time which must actually have elapsed between leaving A and reaching B is $\frac{l\sqrt{1-\frac{v^2}{c^2}}}{c-v}$. Now, we have seen that if we only take account of the slowness of rate of the clock at B, the time elapsed since this clock marked 0 will be $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{l}{c}$. These two quantities are not equal. Hence the clock at B is not merely going slow like the clock at A: it must also be out in its zero. The amount of error is of course $\dfrac{l\sqrt{1-\dfrac{v^2}{c^2}}}{c-v} - \dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}\dfrac{l}{c}$. This comes to $\dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}\dfrac{vl}{c^2}$. Thus when a platform is in motion and clocks are dotted along the direction of its course and synchronised according to the conventions mentioned above, these clocks are not merely all slow in their rate as compared with the same clocks on a platform at rest, but further they are all out in their zeros by an amount that depends on the distance from the standard clock. In fact, to get the "true" time at a place you must not merely take the reading of the clock there and divide it by $\sqrt{1-\dfrac{v^2}{c^2}}$; you must add to the reading an amount $\dfrac{vl}{c^2}$ before dividing by $\sqrt{1-\dfrac{v^2}{c^2}}$. These facts are what are referred to under the name of *Local Time.* It should now be perfectly clear that local time is neither a mystery of the higher mathematics nor a metaphysical whimsy of minds debauched by German philosophy. We have used no mathematics more complicated than simple equations, and we have seen how the experimental facts and the most obvious conventions for synchronising distant clocks force us step by step to this conception.

We have now really got the gist of the famous Lorentz-Einstein transformation. We may sum up our results by saying (i.) that, if you measure a distance l in the direction of motion of a moving platform, the distance between the same two points, if the platform had been at rest, would have been $\dfrac{l}{\sqrt{1-\dfrac{v^2}{c^2}}}$; and (ii.) that if a clock, synchronised according to the method given below with a standard clock at the origin, marks t, the same clock at the same place would have marked $\dfrac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}\left(t+\dfrac{v}{c^2}l\right)$ if the platform had been at rest. The first result is due to the necessity of assuming the Lorentz-Fitzgerald contraction in order to explain the negative result of the Michelsen-Morley experiment. The second is due (*a*) to the general slowing of clocks which has to be assumed to account for the fact that the measured velocity of light is independent of the motion of the platform containing the observers and their instruments, and (*b*) to the variation in the zeros of the clocks which has to be assumed if the velocity of light, as measured by two distant observers on the platform who compare notes, is to be the same as that found by a single observer who stays in one position with a single clock and only notes the phenomena that happen there. (*c*) In the calculations by which we reach the second result we make use of the already assumed Lorentz-Fitzgerald contraction, both in dealing with the rates and in dealing with the zeros of the clocks of the moving platform.

Now, these *results* had already been reached by Lorentz before Einstein came on the scene. The contraction, as its name implies, had been suggested by him and Fitzgerald. The conception of local time and the equation of transformation for time had been introduced by Lorentz for mathematical reasons into which we need not now enter. The originality of Einstein at this stage consisted in his way of connecting these results and viewing them as the consequences of a single general principle. Our next task is to try to understand Einstein's conception of relativity, and the motives for it.

I have stated all the arguments and deduced all the results on the assumption of an ether which is "really" at rest and of clocks which accurately measure the "real" time. I think that there can be little doubt that, with our traditions, this course is *psychologically* the most satisfactory one to follow in order to make the conception of contractions and local time intelligible. But as soon as one reflects on the results one begins to feel that *epistemologically* (to use an unpleasantly pretentious phrase) this whole way of looking at things is artificial and unsatisfactory in the extreme. There are three interconnected points where this becomes specially obvious. (i.) The only velocities for which we have any *direct* evidence are those of one material system relatively to another. Since we cannot perceive the ether we can have no direct knowledge of the velocity of any material system with respect to it. But also—and this has been of the very essence of the business—we have no *indirect* evidence of velocity with respect to the ether; for the main motive for all these transformations has been this very fact, that such supposed velocities have *never* produced the observable effects which they might have been expected to do. In particular, the velocity c of light, though originally defined as its velocity in free ether, has ceased to have any special reference to the ether. For we have seen that the measured velocity of light with respect to platforms moving with *different* velocities is still the same. Now, we *can* tell that two material systems are moving with different velocities, because we can refer them both to a third material system. And it follows that, if there be an ether at all, they must be moving with *different* velocities with respect to it, though we do not know the absolute values of these. Hence we can conclude that, whether there be an ether or not, velocity with respect to it makes no difference to the measured velocity of light or to anything else. The effect of all this is to make the notion of the ether, of uniform motion with respect to it, and of rest in it, utterly unimportant. The results do not, of course, *prove* that there is no such thing; but they do show that, if there be an ether, it is of such a singularly retiring disposition that we need never intrude on its privacy.

(ii.) The Lorentz-Fitzgerald contraction, regarded as a physical phenomenon, is certainly not plausible. In the first place, it has a peculiarity which we shall later on have to notice in connection with gravitation. It is entirely independent of the nature and of the chemical or physical state of the matter concerned. A piece of elastic and a piece of steel would undergo precisely the same contraction if moving with the same velocity. Again, ordinary physical contractions generally have observable physical results. If you strain a piece of glass it begins to exhibit polarisation effects with transmitted light. Such results have been looked for as a consequence of the Lorentz-Fitzgerald contraction, and no trace of them has ever been found. Thus it looks as if this contraction were something quite different from ordinary change of shape and size due to physical stresses.

(iii.) Much the same remarks must be made about the slowing of clocks. It is not easy to see why uniform motion should make all clocks go slower, or why moving a clock in the direction of motion of a platform should upset its zero according to a definite law.

Now, Einstein said: Let us take it as a fundamental principle that uniform motion with respect to the ether makes no difference to the laws of any physical phenomenon. If the ether be a mere fiction this will be necessarily true; and, in any case, so many experiments of the most varied kinds have failed to produce any evidence that such motions are relevant, that it is not rash to generalise their negative results into a principle. Further, let us recognise it as a well-established experimental fact that the velocities of light found by observers in uniform relative motion to each other are the same in spite of their relative motion. Let us then see what follows from these assumptions. Physical laws state relations between the magnitude of some phenomenon in one place and that of some other phenomenon in another. In their most usual form they state that the magnitude of the one is a certain function of that of the other, of the distance between the two places, and of the lapse of time between the two events. So long as the facts are expressed by the *same* function of the same sorts of variable quantities we say that the law remains the same, even though the actual magnitudes of some of the variables should be different. Now, the first part of the principle asserts that if two sets of observers in uniform relative motion observe the same phenomena and discover laws connecting them, these laws will be the same in form. Let the two sets of observers be called A and B, and let them observe the same set of phenomena. It is not a part of Einstein's principle, and it is not in general true, that they will ascribe the *same* magnitudes to the phenomena that they both observe, or regard them as being separated by the *same* distance in space or by the *same* lapse of time. But it is of the essence of Einstein's principle, and it is (so far as we know) true, that the magnitudes that A ascribes to the phenomena, their distances, and the time that elapses between them, will be connected by the same functional relation as the (in general different) magnitudes that B ascribes to the phenomena, their distance, and the time-lapse. *E.g.* if A and his instruments be at rest with respect to an electrically charged body, A's electrical instruments will be giving certain readings dependent on their position with respect to this body, whilst his magnetic instruments will be giving zero readings. If B and his instruments be moving uniformly with respect to this body, both his electrical and his magnetic apparatus will be recording values other than zero. But the law connecting electric readings, magnetic readings, position with respect to the charged body, and time will be precisely the same for both A and B; though of course in the particular function that expresses this law the value that A will put for the magnetic variable will be zero, whilst the value which B will put for the same variable in the same function will be other than zero.

The above may be described as the Physical Principle of Relativity. Taken by itself it would not help us much. What we should like to be able to do would be to pass from a knowledge of the magnitudes that are observable when we are at rest with respect to a system, to those which would be observed if we were moving uniformly with respect to it. For the former are likely to be specially simple, as in the example where the magnetic reading is zero. But a *mere* knowledge that the form of the law must be the same will not enable us to do this. We need also to know the *actual* values that the moving observer will ascribe to at least some of the magnitudes. If we knew this for some of them, in terms of the magnitudes which the resting observer ascribes to the same variables, we could deduce the values ascribed by the moving observer to the remainder. For they will have to be such as to keep the relation connecting all the variables identical for both observers.

This is where the experimental fact of the identity of the measured value of the velocity of light for all observers in uniform motion with respect to each other becomes important. For here we have something that does not merely retain the same *form*, but also retains the same *magnitude* for all such observers. This fact has its implications; we ourselves have seen some of them; Einstein worked them out completely. What he proved was that observers in uniform relative motion will ascribe one and the same velocity to light if and only if certain relations exist between the magnitudes which they each ascribe to the distance between two objects which both perceive. Certain relations must also exist between the magnitudes which each ascribes to the time-lapse between two events which both observe. Now, these relations are precisely those which we have already met with. If one observer reckons a certain distance to be l, another who is moving parallel to this distance with a velocity v as compared with the former will ascribe the length $l\sqrt{1-\frac{v^2}{c^2}}$. If one observer reckons the time when a certain event happens at a certain place to be t, the other will reckon the time when this event happens as $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(t-\frac{v}{c^2}l\right)$, where l is the component, parallel to the direction of relative motion, of the distance between the place where this event happens and the place where the first observer's standard clock is. These are formally the same results as we reached before by considering motion with respect to a supposed ether; but they are now much more plausible and intelligible. Our v was the velocity of the platform with respect to the ether, and our c was the velocity of light with respect to it. But neither of these magnitudes could be measured either directly or indirectly. The v of Einstein's equation is the velocity of any platform with respect to any other, and is therefore directly observable; similarly his c is the velocity of light with respect to any platform, and is therefore again a measurable quantity. We are dealing with nothing but *relative* velocities with respect to *material* systems, and are thus entirely within the region of observable facts. So far, then, from Einstein's way of looking at things being a piece of speculative metaphysics, it is a resolute attempt to be as empirical as possible. It is the consistent application of the principle, enunciated *ad nauseam* by earlier physicists but never really carried to its logical conclusion, that we can and do know nothing but relative motion.

On this interpretation, things which seemed paradoxical and arbitrary on the former method of deducing these transformations become perfectly intelligible. The Lorentz-Fitzgerald contraction ceases to be a physical shortening and becomes a mere question of units of measurement. We have two observers in relative motion. There ceases to be any question of one being "really" at rest and the other "really" moving. Each moves with respect to the other: A counts B to be moving in a certain direction with a velocity v as regards himself; B counts A to be moving in the opposite direction with the same velocity as regards *him*self. Everything is perfectly reciprocal. Each reckons the other man's distances to be shortened and his clocks to be slow by precisely the same amount; and, if we consistently remember the principle that motion is just the rate of change of distance between two pieces of matter, there ceases to be any question of right and wrong between them. Both are right in the sense that they are proceeding on consistent principles and that each will arrive at the same laws of nature.

One question, I think, still remains over. A man might say: "I accept Einstein's physical principle of relativity as both plausible in itself and suggested by the negative results of great numbers of experiments. I also accept it as a fact that observers find the same numerical value for the velocity of light, notwithstanding that they and their instruments are in relative motion. And, since Einstein has proved that this implies certain relations between the measurements of space and the measurements of time used by observers in uniform relative motion, I am logically compelled to accept these relations as a fact. But I do not understand why, as a matter of history, observers in relative motion should have arranged their units of space and time in just this way. As a matter of historical fact, they must have set up their conventions for measuring distances and time-lapses without a thought of the velocity of light. Since they do find afterwards the same numerical value for this, their measures of distance and of time-lapse must in *fact* have had the relations which Einstein asserts; but I do not in the least see why they should have done so. Their choice of units was in their own power; they made their selection without any reference to the velocity of light; surely it is an extraordinary coincidence that the units which they actually hit upon should have happened to stand in these relations."

This is a perfectly reasonable question to raise. It can be answered, I think, by reflecting on the way in which our judgments about identity of length, sameness of rate, and identity of time begin and develop. We start with crude immediate judgments on such matters. As our researches become more accurate we develop new and more searching tests for congruence, isochronism, simultaneity, etc. But these tests always contain an element of convention; and the more minute the differences with which we are dealing, the bigger will be the dose of convention. Let me explain. The first step is to put two rods or two clocks in such positions that differences of length or in the time of swing of pendulums can be noted with special ease if they exist. So far there is but little convention present; we are still resting on our immediate judgment, and are simply arranging objects in such a way that such judgments shall have the chance of being as accurate as possible. But, when we pass beyond this point, perception and immediate judgments of congruence have done all that they can do. Our further refinements, our more accurate tests, must be of a different nature. We are now, by hypothesis, dealing with differences too small to be *directly* perceived even under the most favourable circumstances. Hence our tests must now involve the supposed perceptibly different *consequences* of *imperceptible* differences of magnitude. They thus imply that there are certain laws of nature, relating imperceptible differences in the magnitudes to be compared and perceptible differences in something else, and connecting the two, moreover, by some mathematical relation which will enable us to infer the size of the former differences from that of the latter. All such laws are necessarily hypothetical, since, by hypothesis, one term in the relation (the imperceptible differences of magnitude) can never have fallen under direct observation. The assumed laws are therefore largely in our own power, and, according to the special form that we suppose them to have, the magnitudes ascribed to imperceptible differences will differ, even while the resulting differences in the observable magnitudes remain the same. Thus our more accurate and minute measures of magnitude, judgments of congruence, etc., depend upon the special form that we choose to give to laws which are necessarily hypothetical. We naturally try to make these laws as simple as possible, and also as much in accord as possible with what we have gradually learned about the general "make-up" of the material universe. This very greatly restricts our choice. Now, all the differences between Einstein and commonsense are extremely small, depending as they do on terms of the order $\frac{v^2}{c^2}$, where v is the velocity of one piece of matter relatively to another, and c is the enormously greater velocity of light. But only a very small group of men have ever had to deal with measurements carried to this order of accuracy. And all these men have been physicists, saturated with a common tradition, and holding substantially the same view as to the general "make-up" of nature. It is therefore not in the least surprising that they should all have hit on the same conventions as to the measurement of space and time. Some of these were used quite unconsciously, and it is a great merit of Einstein to have dragged them to the light and deduced their consequences. I think that on these lines a satisfactory answer can be made to the question which I supposed an intelligent objector to raise. If the reader wants an illustration, he has only to refer to our previous discussion of the synchronising of distant clocks. The criteria there used no doubt seem a little odd so long as we assume an ether and suppose a platform moving through it. We might be inclined to say: Surely they can see that their criterion for identity of zero, *e.g.*, is only valid if the platform be at rest. But, once we clearly understand that motion or rest with respect to the ether cannot be detected and may be the merest fiction, it becomes clear that this is the only criterion that the observers can use, and that it is as reasonable for those on one platform to use it as for those on any other; for there is no sense in saying that one is at rest and the other is moving. Each is at rest with respect to itself, each is moving with respect to the other, and these are the only motions that we can detect and deal with.

When once the transformation equations for space and time measurements have thus been established, the physical principle of relativity comes into play. We find that a certain law connects a set of phenomena, their distance, and the time-lapse between them, when all these are measured by an observer at rest with respect to these phenomena. We ask: What magnitudes will be ascribed to these phenomena by an observer who is moving uniformly with his instruments with respect to the first? The physical principle tells us that the *form* of the law will be the same for both, and we know what that form is for the first observer. The transformation equations tell us the values that the second observer will ascribe to the distances and time-lapses in terms of the values which the first ascribes to them. We have therefore merely to see how the magnitudes ascribed by the second observer to the phenomena must be related to those which the first observer ascribes to them in order that the form of the law may be the same for both.

The last point to notice is this. Many physical laws were already in such a form that they accorded with the principle of relativity. Examples are Maxwell's equations for the electromagnetic field, and the equation of continuity in hydrodynamics. But other laws, as stated, were not in accord with the principle. As the principle is perfectly general, such supposed laws needed slight modification to make them admissible laws of nature. *E.g.* it is incompatible with the principle of relativity to hold both that momentum is conserved and that mass is wholly independent of velocity. If we keep the former belief we must suppose that the mass of a particle, as reckoned by an observer moving relatively to it, differs from the mass of the same particle as measured by one who moves with it. The difference depends on terms of the order $\frac{v^2}{c^2}$. Now, this had already been predicted and verified for electrons shot out with great velocities in vacuum tubes. In fact, the greatest triumph of the restricted theory of relativity has been that numbers of results, which had formerly been predicted from special *physical* hypotheses and verified, simply tumble out as consequences of the principle without needing any special physical hypothesis at all.

I trust that I have now clearly explained Einstein's restricted theory of relativity, its grounds, and its consequences. We are now in a position to try to understand his generalised theory, which involves the new views about gravitation. In what sense is the principle of relativity sketched above "restricted"? It is restricted in the sense that it only refers to motions which are rectilinear in direction and constant in magnitude. It does not follow from the restricted theory that if one observer and his instruments be accelerated with respect to another, or be rotating about the other, the form of the laws of nature will remain the same for both. Yet of course, if we are to be in earnest with the view that all motion is relative, that it is always simply a change in the respective positions of material systems, accelerations and rotations are just as relative as uniform rectilinear translations.

Now, long before Einstein, indeed ever since Newton, accelerations and rotations have been a stumbling-block for a purely relative theory of time and space. Newton's laws of motion (the third law in particular) assume "unaccelerated axes." Given a set of axes such that motions with respect to them obey Newton's laws, any other set of axes that moves with a uniform translatory motion with respect to these will do equally well. But, if you refer the motions of a system to axes *accelerated* with respect to the set mentioned above, these motions will not be subject to Newton's laws. Thus a certain group of sets of axes which we will call *Newtonian*, and which is such that any pair of sets from the group are either relatively at rest or relatively in uniform motion in a straight line, occupies a privileged position. But not every group obeying these conditions will be Newtonian; and you must either *define* the Newtonian group by the fact that motions with respect to any member of it obey Newton's laws, or by the fact that all members of it are unaccelerated. The former procedure implies that Newton's laws are not true for *all* sets of axes, and thus prevents us from applying a generalised principle of relativity to them as they stand. The latter suffers from the two defects that it assumes absolute space, time, and motion—as of course Newton did,—and that, whether there be such things or not, they cannot be perceived, and therefore cannot be the actual criterion by which we in fact determine whether a set of axes is or is not Newtonian.

The difficulty about rotation is excellently illustrated by an example of Einstein's which I shall give in my own words. (It had, of course, already been discussed by Newton in his famous "bucket experiment," and used by him as a proof of the necessity of distinguishing between absolute and relative rotation.) Suppose there were two masses of liquid so far from each other and from other matter that each is subject to no force except the gravitational attraction among its own particles. Each will assume a spheroidal form under the action of these internal forces. Now, suppose that on the surface of each (say on the equator) there is a mark that can be seen through a telescope by the inhabitants of the other. Lastly, suppose that the inhabitants can communicate with each other by wireless, and that each can carry out a survey of the surface of his spheroid. Suppose that the people on A noticed that the mark on B was rotating with a velocity Ω about the common axis of symmetry. Then of course the people on B would equally judge that the mark on A was rotating with the same velocity about the same line in the opposite sense. This information they could communicate to each other. But if they then proceeded to survey their respective spheroids, A might be found to be a perfect sphere and B to be flattened at the poles, *i.e.* at the points where the common axis of symmetry cuts the surface. When this information was communicated difficulties would arise. Each is rotating in precisely the same way as judged from the other; why then should one be flattened and the other remain spherical? *We* might say: "B is 'really' rotating, and A is 'really' at rest, and that is the cause of the difference." But this assumes absolute rotation; if we confine ourselves to relative rotation, the circumstances of each are precisely similar, and the observable difference of shape is a mystery.

This is Einstein's example, and it is such a striking one that it will be worth our while to discuss it much more fully than he does. Suppose we agree to drop the notion of absolute rotation, what alternatives are open to us? (i.) Mach would say that the whole example is like discussing whether beggars would ride astride or on side saddles if wishes were horses. We only know how matter behaves in the presence of the whole stellar universe; if that were away, as the example assumes, both masses of liquid might be spherical or both flattened. Either we assume that the two masses of liquid actually are the only matter in the universe, or that the fixed stars, though assumed to be too distant to produce appreciable gravitational effects or to be seen, still exist. On the former alternative we have not the least idea what would happen, because the conditions are so widely different from those under which all our actual observations have been made. On the latter we shall simply have to say that B is rotating with respect to the fixed stars and A is not, and that this difference is the source of the differences which only puzzle the inhabitants because their eyesight is not strong enough to see the fixed stars. The first part of Mach's objection seems to me obviously sound; we really have no right to conjecture what Einstein's two masses of liquid would do if they constituted the whole material universe. We must therefore assume that the material universe as a whole is much as we know it, and that the two liquid blobs are simply very remote from other bits of matter which in fact exist. On this hypothesis we may generalise the present answer to the difficulty by saying that there is ordinary matter concealed from the inhabitants of the spheroids; that in fact one of them is rotating with respect to the concealed matter, and the other is not; and that these relative rotations which they do *not* notice have physical consequences which do not follow from the relative rotations that they do notice. This is a logically possible view. Einstein goes further and tries to prove by an epistemological argument that it is the *only* possible one ("Die Grundlagen der allgemeinen Relativitätstheorie," *Annalen der Physik*, 1916, No. 7). And he deduces from it that "the laws of physics must be so constituted as to hold good in relation to any set of axes, however it may be moving." As regards the first point his argument simply is that the law of causation is a law of phenomena, and therefore any epistemologically acceptable cause for the observed flattening of B must be something that is in principle capable of being perceived. It must therefore be something that happens in ordinary matter, and not anything that involves absolute space or ether. This argument he calls "*schwerwiegend*"; it seems to me to rest on the merest prejudice, inherited probably from Kant. An electron is apparently a permissible cause, the ether is not. But in actual fact you can perceive neither. If you say that we could perceive an electron provided certain alterations were made in our senses, it is always open to an opponent to say that, if there be any ether at all, we could perceive *it* provided suitable modifications were made in our senses. The fact is that anything that could exist could in theory be perceived if we had the right kind of senses, and the question whether our senses would need much or little modification in order to perceive a suggested entity has not the least bearing on the question whether that entity exists and can be taken as a *vera causa.* Thus no weight can be laid on Einstein's epistemological argument to prove that the explanation of the flattening of B *must* be found in its relation to other parts of the *material* world which are imperceptible from A and B.

(ii.) Thus an alternative is open to us, which Einstein erroneously believes to be cut out by his epistemological argument. Why should we not say that here at last the ether has emerged from its *otium cum dignitate* and produced a measurable effect on ordinary matter? Why not say, in fact, that B is rotating with respect to the ether, and that such rotations cause flattening; whilst A is at rest with respect to the ether, and therefore remains spherical? There would be no contradiction in this to the restricted principle of relativity, for that did not disprove the existence of the ether, and only asserted that uniform translation with respect to it (if it existed) made no difference to any observable phenomenon. Nor, again, should we be giving up the view that all motion is relative, though we should of course be dropping the view that it is always relative to ordinary matter. There are even some positive advantages in this as compared with the Mach-Einstein view. I have said that the interpretation put by Einstein on his example is logically possible though not epistemologically necessary. But it has its difficulties. Its great defect is that it puts some pieces of matter in an unintelligibly privileged position in the universe. Rotations with respect to these pieces of matter (*e.g.* the fixed stars) have physical effects; rotations of precisely the same kind with respect to other pieces of matter (*e.g.* those of the spheroid A with respect to the spheroid B) have no such consequences. Now, there is nothing mysterious or unusual about the fixed stars. It is extremely difficult to see why certain perfectly ordinary bits of matter, distinguished by nothing else from other bits, should stand in this exceptional position, especially when one remembers that, in the case of the fixed stars at least, their one outstanding feature is their extreme *remoteness*, which is the last factor that one would expect to be associated with special causal efficacy. The present alternative avoids this difficulty; the rotations which have special importance are rotations with respect to a special kind of matter (viz. the ether).

Is Einstein justified in concluding that, if all motions be those of one piece of ordinary matter with respect to another, the laws of physics must be capable of statement in a form independent of the movements of our axes? He just states this dogmatically; but I imagine that, when expounded, the argument would run as follows:—Suppose you choose any one set of material axes and any one periodic process as a time-measurer. The motions of *all* other pieces of matter with respect to these axes and this time-measurer will obey certain laws, which will be simple or complex according to the bodies that you have chosen as axes and the process that you have chosen to measure time. Now, suppose you choose any other set of bodies for axes and any other set of events for time-measurers. The motions of any material system with respect to the new axes can be compounded out of their motions with respect to the old axes and the motions of the old axes with respect to the new ones. Now, the laws of the former are known. And the latter are uniquely connected with the motions of the new axes with respect to the old ones. And the laws of these are known. Hence the laws of motion with respect to the new axes are simply a mathematical transformation of the laws with respect to the old axes. I suspect that this is Einstein's meaning, and, if so, he seems to me to be right.

The upshot of the discussion so far seems to be this:—Einstein is mistaken in thinking that he can prove epistemologically that all motion must be that of one piece of ordinary matter with respect to another. But if this be in fact true, he is right in supposing that it follows that the laws of physics ought to be capable of statement in a form that is independent of our particular choice of spatial axes and temporal rate-measurers. And it will evidently be a great triumph if we can succeed in doing this; it will be at once a great extension of the principle of relativity, and a setting to rest of all the difficulties that have sprung from the fact that, ever since Newton stated his laws of motion, we have wanted to believe that all motion is that of one piece of matter with respect to another, and have at the same time felt that these laws involved (as their author asserted) a distinction between absolute and relative motion.

Let us now consider the application of the principle a little more in detail. Newton's first two laws say that a material particle is either at rest or in uniform rectilinear motion when not acted upon by forces, and that if a force acts upon it the particle will be accelerated in the direction of the forces by an amount equal to the force divided by the mass. Now, if we drop absolute space, time, and motion, all these statements are vague in the extreme. A particle at rest with respect to one set of axes will move in a straight line with respect to another set, and in some other curve with respect to a third set. And if time be measured by one process it will be moving uniformly, whilst if another process be used it will be accelerated. *E.g.*, suppose we have a particle at rest relative to a platform on which it is lying. Let there be a wheel also lying on this platform, and let us take two mutually rectangular spokes of it as our axes. Then, so long as the wheel remains still with respect to the platform, the particle will be at rest with respect to these axes and will therefore be said to be under the action of no forces. But suppose that the wheel begins to rotate with respect to the platform with uniform angular velocity Ω. With respect to the axes the particle will now describe a circle about the centre of the wheel with this angular velocity. If all motion be purely relative this rotation with respect to one set of material axes is as genuine a fact as its rest with respect to the other. But a rotation with angular velocity Ω implies an acceleration towards the centre of amount $r\Omega^2$, where r is the distance from the centre of the wheel to the particle. People who use this set of axes will therefore say, in accordance with Newton's laws, that the particle is acted upon by a force $mr\Omega^2$ towards the origin. It is therefore perfectly open to us to choose any axes we like and to hold that Newton's first two laws apply to all motions with respect to them; but we shall have to assume in general that different forces are acting according as we refer the motion to one set of axes and one rate-measurer or to another. Thus force ceases to be something given once and for all; the field of force that has to be assumed, if Newton's first two laws are to hold independently of choice of axes, itself depends on the axes chosen and the process of time-measurement used.

It will be noticed that I have carefully confined the above assertion to Newton's first two laws. The reason is this. Newton's third law implies that force *on* one particle is always a one-sided way of looking at an event which in reality consists of a stress *between* two particles. This means that when we find a force acting on a particle we must always expect to find that this is due to some other particle on which the first exercises an equal and opposite force. Now, if we call forces that obey this law Newtonian forces, we shall have to admit that some at any rate of the forces that are connected with a change of axes are non-Newtonian. Let us revert to our example of the particle and the wheel. The people who choose two spokes of the rotating wheel as their axes can, as we saw, keep Newton's first two laws, provided that they introduce with their new axes a new force of amount $mr\Omega^2$ acting on the particle towards the centre of the wheel. But this force is non-Newtonian; there is no equal and opposite force on the centre of the wheel to balance it, as there should be by Newton's third law. If there were, the centre of the wheel should either be accelerated or at least subject to a pull. It is clearly not accelerated with respect to the axes under discussion, for it is and remains their origin. And, although it is no doubt being pulled by the spokes, the magnitude of the pull has no connection with the mass of the particle, but only with the inertia of the wheel. Thus, to keep Newton's first two laws true when motion is referred to the new axes, we have had to introduce a force that breaks his third law. Such forces have another peculiarity. They are independent of the physical or chemical state of the bodies on which they are exerted and of the medium in which they may be swimming. Every particle, whatever it may be made of and whatever may surround it, will be accelerated by an amount $r\Omega^2$ with respect to these axes, and therefore the force that will have to be assumed to keep Newton's first two laws true will depend on no property of the particle except its mass.

Now, Einstein observes that the force of gravitation stands out from all other physical forces by possessing just the peculiarities that we have noted for non-Newtonian forces. Particles in a gravitational field are acted upon by forces that depend on no property of the particle except its mass. Most elaborate experiments have been performed to test whether the acceleration produced by a gravitational field on bodies depends in any way on their temperature, crystalline form, chemical composition, surrounding medium, etc. No trace of any such dependence has been found; the one relevant factor seems to be their mass. Thus there is a very strong motive for treating gravitation as one of those non-Newtonian forces that are associated with changes of axes or of time-measurer.

Just as we introduce forces which were not present with one set of axes by passing to a new set, so, in many cases, we can get rid of forces which were present with one set by using another. If we are dealing with phenomena in space just over Trafalgar Square, and use Nelson's column and the lines joining the diagonally opposite lions as our axes, our phenomena will be taking place in a practically constant gravitational field. If we drop a stone from the top of the column and take as our new axes lines in this stone parallel to our original axes, we have "transformed away" the gravitational field for this region. *I.e.* if we refer the same phenomena to the new axes, we can treat them as not being under the action of gravitation. For any bodies that are falling to the earth will now be unaccelerated with respect to our axes. Other bodies, such as Nelson's statue, which are not falling, have an upward acceleration with respect to our new axes, and are therefore subject to a force which, in the particular example, we should ascribe to the reaction of the column on the statue. (With the old axes this reaction was also present, but balanced by the gravitational pull on the statue.)

It is not, however, in general true tha gravitational forces can be wholly transformed away by suitable choice of axes over every finite region; for the field in such a region is not in general the same everywhere or always. It is greater in parts of the region that are near large pieces of matter, and it may change with time as bodies approach the region. Thus we cannot guarantee that any change of axes will get rid of it permanently and for the whole region. What we can say, however, is this. The smaller we make the region, the more nearly can we transform away the gravitational forces by suitable choice of axes; and we assume that in an infinitesimally small region for an infinitesimally small time a set of axes and a time-measurer can always be found such that with respect to these there are no gravitational forces.

Let us call a set of axes and of clocks such that the restricted principle of relativity holds with respect to them a Newtonian frame of reference. The principle does not hold with respect to any and every frame of reference. *E.g.*, with respect to our wheel that rotates relatively to a Newtonian frame light evidently does not travel in straight lines or with a constant velocity. Thus the restricted principle, which assumes the rectilinear propagation and the constant velocity of light, presupposes a Newtonian frame of reference.

Now, if we could find some magnitude connected with a pair of events closely adjacent in space and time, and such that it was independent of the system of axes and the time-measurer chosen, we should have found a certain common condition that all possible frames of reference must obey. There *is* such a magnitude for any adjacent pair of events; it is called their "separation," and is denoted by the symbol $d\sigma$. To get some idea of the notion of separation let us consider two adjacent points in space. They determine an unique magnitude, their shortest distance—usually denoted by the symbol ds,—and this is the same whatever spatial axes we adopt for giving co-ordinates to the two points. Now, we have to deal, not with adjacent *points*, but with adjacent *events*. Hence we want an extension of the notion of the distance between two points which shall include also the time-lapse between two events. As I have said, there is such a magnitude, the separation $d\sigma$; and separation is best regarded as an extension of the notion of distance.

Now, we know how the separation is connected with the co-ordinates and the time for a Newtonian frame of reference. Suppose that two events happen at two adjacent points whose spatial co-ordinates with respect to Newtonian axes are respectively x, y, z and $x+dx$, $y+dy$, $z+dz$. Suppose further that the first event happens at a time t, and the second at a closely adjacent moment $t+dt$, as measured by a Newtonian clock. Then the separation is given in terms of the co-ordinates and of the time by the expression

$$d\sigma^2 = -dx^2 - dy^2 - dz^2 + c^2dt^2.$$

This fact is a consequence of and is equivalent to the fact that the velocity of light with reference to all Newtonian frames is c.

Now, the fact that $d\sigma^2$ depends solely on the two events and not on the particular frame chosen for placing and dating them imposes a condition upon all possible frames of reference. Every possible frame will be characterised by the equation that expresses the separation in terms of the co-ordinate differences of the two events with respect to the axes and the time-lapse between them as measured by the clocks of the frame. Thus, whatever frame you choose, there will be a certain function of the co-ordinates and the times of the two events with respect to it which must be equal to the function $-dx^2 - dy^2 - dz^2 + c^2dt^2$ of the corresponding Newtonian magnitudes for the same pair of events.

Thus, although an infinite choice of frames of reference is open to us, it is not absolutely indefinite. Nature imposes a certain very general restriction upon all frames of reference that can be used for dealing with natural phenomena. For events that take place at an infinite distance from matter the expression for the separation in terms of the co-ordinates and the time will take the specially simple form contemplated by the restricted theory. But everywhere else there will be gravitational forces; and if you choose such a frame as will transform them away over a small region, which we have seen you always can do, the expression connecting the separation with the new co-ordinates and the new time will be different. Thus the common condition imposed on all possible frames expresses the universality of the law of gravitation, and the particular form of the expression for the separation expresses the special gravitational field in the small region for which this particular frame has to be used if the field in that region is to be transformed away.

Even apart from these special considerations we can see in a general way that the law of gravitation imposes certain limitations on our frames of reference. Gravitation is supposed to act between all pieces of matter in the universe. If we are in earnest with the view that all motion is the change of position of one piece of matter with respect to others, all sets of axes that we can possibly choose will be material, *i.e.* they will be defined by certain actually existing pieces of matter. Thus all possible axes will themselves exert some gravitational attraction on every body in the universe, and therefore on every body whose motion is referred to them. Thus the gravitational attraction between axes and what is referred to them is a feature common to all possible frames of reference; the most we can say is that the further the referred bodies are from the axes of reference, the less this influence will be. Thus we see that the statement that all possible frames of reference are subject to a certain limiting condition, and that this condition embodies the law of gravitation, is not a wild paradox which we can only accept through the *force majeure* of a "knock-down" mathematical proof. It is a fact which commonsense and our previous ideas about the universality of gravitation might have suggested to us; so that we can regard the mathematical arguments rather as clearing up the details of what was previously a vague general anticipation than as ramming a new and utterly unforeseeable fact down our throats.

I hope that I have now succeeded in giving the reader at least a rough idea of the meaning and the motives of Einstein's theory of gravitation. It will be noticed that I have done so without saying a word about non-Euclidean geometry. This seems to me to be an advantage in a statement of the theory to persons who are unfamiliar with the concepts of that branch of mathematics. But my paper would be incomplete if I left matters at this stage. Everyone has heard that the new theory has a great deal to say about Euclidean and non-Euclidean space, and I shall conclude by trying to indicate the connection of the two subjects.

For this purpose it will be best to return for the moment to the restricted theory of relativity and to explain Minkowski's geometrical representation of it. Everyone has seen a recording barometer. In this instrument a drum covered with paper rotates at a uniform rate whilst a pen-point attached to the barometer presses against it. As the barometric pressure rises and falls, so does the pen. When the paper is unrolled at the end of the week we find a curve on it. One axis represents the time elapsed, the other the position of the top of the mercury column. Thus any point on the curves represents a momentary event, viz. the presence of the top of the mercury column at such and such a height at such and such a time. Now, the notion of Space-Time[1] is simply an

[1] I borrow this convenient name from Prof. Alexander without necessarily using it in the same sense as he does.

extension of this. Every event in the world happens somewhere and somewhen. To define its spatial position we need three spatial co-ordinates, to define its temporal position we need one temporal co-ordinate. Thus to represent events in general in the same kind of way in which the particular event which is the momentary position of the top of the mercury column is represented we need a four-dimensional diagram. Naturally we cannot draw such a thing, but we can treat it analytically just as easily as one of three or of two dimensions.

Suppose that we choose as our axes for space-time the ordinary x, y, z of a Newtonian frame, and ct, *i.e.* the time measured by a Newtonian clock multiplied by the velocity of light. Any event that ever happens will be represented by a point in space-time. The history of a particle that is moving about will be represented by a curve in space-time, and, if the particle happens to be moving in a straight line with uniform velocity, this curve will be a straight line. But, if our space-time is to do the work required of it, it must differ from an ordinary Euclidean space, not merely in the fact that it has four dimensions instead of three (a comparatively trivial distinction), but also in the fact that the "distance" or "separation" between two adjacent points in it is related to the differences of their co-ordinates in another way than it would be in a Euclidean space. In an ordinary three-dimensional Euclidean space the distance between two adjacent points is related to the differences of their co-ordinates by the equation

$$ds^2 = dx^2 + dy^2 + dz^2 \quad . \quad . \quad . \quad \text{(i.)}$$

If, therefore, space-time were Euclidean, though four-dimensional, the separation between two points in it would be given by

$$d\sigma^2 = dx^2 + dy^2 + dz^2 + c^2dt^2 \quad . \quad . \quad . \quad \text{(ii.)}$$

But actually the restricted principle of relativity requires that the relation should be

$$d\sigma^2 = dx^2 + dy^2 + dz^2 - c^2dt^2 \quad . \quad . \quad . \quad \text{(iii.)}$$

Now, this sort of relation is characteristic of an hyperbolic space, *i.e.* of the sort of space first noticed and discussed by Lobatchewsky. It is easy to see in a rough general way why the restricted theory requires the relation (iii.) instead of the relation (ii.). Suppose that one event is the fact that a beam of light has reached a point x, y, z at a moment t, and that the adjacent event is that the same beam has reached the point $x+dx$, $y+dy$, $z+dz$, at $t+dt$. The distance travelled by the beam in space is ds, which by (i.) is equal to $\sqrt{dx^2+dy^2+dz^2}$. The time taken between the two events is dt. Since c is the velocity of light, and one event is the arrival of light at one place and the other is the arrival of the same light at the second place, we must have $ds=cdt$.

Thus for this pair of events the formula (ii.) would give

$$d\sigma^2 = c^2dt^2 + c^2dt^2 = 2c^2dt^2,$$

whilst the formula (iii.) would give

$$d\sigma^2 = c^2dt^2 - c^2dt^2 = 0.$$

Now, on the restricted theory nothing can move faster than light. Thus we should expect the separation between two events which are respectively the departure and the arrival of the same beam of light at two adjacent places to be the minimum possible separation. This is secured by the relation (iii.) but not by the relation (ii.). So the space-time of the restricted theory of relativity is not Euclidean but is hyperbolic.

Now, the transformations of the restricted theory can be put in a very striking form when stated in terms of this non-Euclidean space-time. It can be shown that the relations which we have seen to exist, on the restricted theory, between the values which two observers in uniform relative motion ascribe to the co-ordinates and the time of the same event, can be interpreted as follows:—You have simply to imagine the axes of your space-time twisted, without changing the angles between them, by a certain amount about their origin; the co-ordinates of a given event with respect to the new axes will then be related to the co-ordinates of the same event with respect to the old axes by precisely those relations which we deduced from the restricted theory of relativity. We may put it in this way. The change involved in your spatial co-ordinates and your measure of time, when you pass from one platform to another which moves uniformly with respect to it, is completely represented by twisting the axes of Minkowski's hyperbolic space-time about their origin, without change of their mutual relations, through a suitable angle. Now, everyone admits that if you take a set of three *spatial* axes at right angles to each other and intersecting at a point, it will make no difference to the laws of nature to twist them as a rigid body about this common point. Suppose, *e.g.*, that the spatial axes were the three edges of a biscuit-tin that meet in a corner: you could evidently turn the tin into any position about this corner without altering the form of the laws of nature. The restricted principle of relativity is equivalent to a generalisation of this fact, which is so obvious for space, so that it also includes time. For it tells us that the form of the laws of nature is unaffected by passing from one platform to another in uniform motion relative to the first. And this change we have seen is equivalent to a twist of axes in *space-time*, comparable to the twisting of our biscuit-box about one of its corners in *space*. The analogy is not absolutely complete, because space-time is hyperbolic, whilst the space of which the edges of the biscuit-tin form a set of axes is Euclidean. And a twist of a body in hyperbolic space is not quite the same thing as the twist of one in Euclidean space. Still, the analogy is great enough to render this a most striking and helpful way of visualising the restricted principle of relativity. There is no need, so far as I can see, to suppose that this representation is anything more than an attractive mathematical device; it no doubt *invites* us to develop metaphysical theories about space-time, and this is worth doing when it is undertaken by competent people like Mr Robb, Dr Whitehead, and Prof. Alexander. But it certainly does not *necessitate* anything of the kind, and philosophers in general will be unwise to rush in where physicists fear to tread.

We ought now to have little difficulty in understanding how the new theory of relativity is connected with non-Euclidean geometry. A Newtonian frame of reference is one for which the restricted principle of relativity holds, and is therefore represented by the hyperbolic space-time of Minkowski, which we have just been discussing. Such a space-time is not indeed Euclidean, but it shares with Euclidean space the important property of being "homaloidal." Roughly speaking, this means that it is everywhere alike: the separation of two points depends only on the differences of their co-ordinates. This is obvious from the expression

$$d\sigma^2 = dx^2 + dy^2 + dz^2 - c^2dt^2,$$

where nothing occurs on the right-hand side except the differences of the co-ordinates of the two points in space-time, and certain *constants* which multiply these, viz. 1, 1, 1, and $-c^2$. But we have also seen that Newtonian frames can only be found for finite regions at places infinitely remote from matter. Near gravitating bodies there are forces; we have seen that for infinitesimal regions these forces can be "transformed away" by a suitable choice of axes, but these axes will have to move in various complicated ways with respect to Newtonian axes in order to compensate for the acceleration which the gravitational attraction produces with respect to Newtonian axes. Such moving axes will obviously be represented by a very different kind of space-time from that which represents a Newtonian frame. The nature and the motion of the frame will be completely represented by the expression for ds^2 in terms of the differences of the co-ordinates and of the time of the frame. In general we shall get an expression for the separation of the form

$$ds^2 = g_{11}dx'^2 + g_{22}dy'^2 + g_{33}dz'^2 + g_{44}dt'^2 + g_{12}dx'dy' + g_{13}dx'dz' + g_{14}dx'dt' + g_{23}dy'dz' + g_{24}dy'dt' + g_{34}dz'dt',$$

very different from the elegant simplicity of the Newtonian frame. Now, these g's will not in general be constants: they will themselves be functions of x', y', r', and t'. Thus the geometry of space-time in such a part of the universe will differ wildly from the Euclidean and even from the tamer kinds of non-Euclidean geometry to which we have become accustomed.

The position, therefore, is this:—In the neighbourhood of a piece of matter (*e.g.* the sun) all bodies are acted on by gravitational forces with respect to a Newtonian frame. You can transform away these forces for small regions by a suitable choice of moving axes. The particular frame needed for this purpose can be expressed in terms of the geometry of space-time for the region, just as Minkowski expressed the Newtonian frame by a hyperbolic space-time. The complete geometry of space-time for the region is summed up in the form of the g's which appear in the expression for the separation. These g's can therefore be regarded from two points of view. (i.) From one point of view the g's express the forces that bodies in this region would experience, judged from a Newtonian frame. (ii.) From the other point of view the g's express the geometry of space-time for that particular non-Newtonian frame of reference with respect to which these forces have been transformed away. This geometry is in general wildly non-Euclidean.

I will conclude by stating the practical consequence, for physical purposes, of the new theory of relativity. The extended principle is that any genuine law of nature must have a form independent of the frame of reference that we happen to use for placing and dating phenomena. This by itself would not be of much value unless there is something that keeps not merely its *form* but also its *value* fixed for all possible frames. We noticed just the same fact about the restricted theory: there the constancy of the velocity of light for all Newtonian frames came to our help; here the constancy of the separation of two adjacent events for all frames whatever plays the same part. Now, we have a good many laws of nature already stated with respect to Newtonian frames, *e.g.* Maxwell's equations. We now know that they must be capable of statement in a form that is independent of any particular frame. It is therefore our task to find this form, guided by the two facts (*a*) that we know the form of the law for the specially simple case of a Newtonian frame, and (*b*) that we know that any possible frame must be so related to a Newtonian one that the value of the separation of the same pair of adjacent events is the same for both. With these facts it is possible to solve the problem by means of a certain branch of pure mathematics called the Absolute Differential Calculus, which had been developed for other purposes by Riemann, Christoffel, and Levi-Civita. As with the restricted theory, we find that some laws have already been stated in a form consistent with the principle of relativity; others have not. As before, Maxwell's equations obey the principle without any modification; Newton's law of gravitation does not, but needs a modification which makes a difference that is excessively small in all but a few cases. One of these cases is the position in space of the perihelion of Mercury: on the old law it should be fixed; on the law as modified to meet the principle of relativity it should gradually change its position. This it actually does, and by almost exactly the amount predicted by the new theory.

Finally, we must notice the following important consequence of the theory. We have seen that for small regions a frame can always be chosen that will transform away the gravitational forces. Thus for a sufficiently small region the presence or absence of a gravitational field is simply equivalent to the use of one or another frame of reference. But the form of the laws of nature is independent of the frame of reference chosen. Therefore for a sufficiently small region the form of the laws of nature should be independent of the presence or absence of a gravitational field. The *amount* of independence will depend on the size of the region for which the field can be transformed away by a mere change in our frame of reference. Thus we might expect that some laws will change their form in a gravitational field and that others will not. Now, light moves with an uniform velocity in a straight line with respect to a Newtonian frame when there are no forces. This is the particular case for Newtonian frames of a law of the general form that light travels in such a path between two points as to make the time-lapse a minimum. This law holds for all possible frames of reference. Now, the gravitational attraction near the sun may be transformed away by choosing a suitable frame, which will of course be non-Newtonian. So light there will, with respect to this frame, be under the action of no forces, and will therefore move so that the time-lapse is a minimum. But a path which, in respect to the new non-Newtonian frame, fulfils these conditions will not do so with respect to a Newtonian frame. Thus, judged from a Newtonian frame, light that passes near the sun will not move with an uniform velocity nor in a straight line. The deflection can be calculated on Einstein's theory, and it has been verified by observation.

I have now fulfilled my promise to the best of my ability. We have seen what exactly Einstein's theory is and how it is related to Euclidean geometry and to Newtonian mechanics. The connection with the former is not really very intimate, and Einstein himself makes very little play with it. The connection with the latter is all-important. Einstein's discovery synthesises Newton's two great principles—the laws of motion and the law of gravitation. It removes the obscurity that has always hung over the former, by working out the relativity of motion to the bitter end, whilst it generalises and slightly corrects the latter and accounts for its peculiar position among all the other laws of nature. Such work can only be done by a man of the highest scientific genius, and we have no right and no need to enhance his greatness by decrying the immortal achievements of his predecessors. It is enough that we can, without the slightest flattery or hyperbole, class Einstein with Newton, and say of the former what is written on the tomb of the latter:—

"*Sibi gratulentur homines tale tantumque exstitisse humani generis decus.*"

C. D. BROAD.

University of St Andrews.