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

$$\frac{mc^2 + E_0}{\sqrt{1 - \frac{v^2}{c^2}}},$$

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

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

Let me add a final remark of a fundamental nature. The success of the Faraday-Maxwell

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

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

XVI

EXPERIENCE AND THE SPECIAL THEORY OF RELATIVITY

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

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

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

In the theoretical treatment of these electrons, we are faced with the difficulty that electrodynamic theory of itself is unable to give an account of their nature. For since electrical masses of one sign repel each other, the negative electrical masses constituting the electron would necessarily be scattered under the influence of their mutual repulsions, unless there are forces of another kind operating between them, the nature of which has hitherto remained obscure to us.[1] If we now assume that the relative distances between the electrical masses constituting the electron remain unchanged during the motion of the electron (rigid connection in the sense of classical mechanics), we arrive at a law of motion of the electron which does not agree with experience. Guided by purely formal points of view, H. A. Lorentz was the first to introduce the hypothesis that the particles constituting the electron experience a contraction in the direction of motion in consequence of that motion, the amount of this contraction being proportional to the expression $\sqrt{1-\frac{v^2}{c^2}}$.* This hypothesis, which is not justifiable by any electrodynamical facts, supplies us then with that particular law of motion which has been confirmed with great precision in recent years.

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

[* $\sqrt{1=\frac{v^2}{c^2}}$ — J.M.]

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

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

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

XVII

MINKOWSKI'S FOUR–DIMENSIONAL SPACE

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

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

Similarly, the world of physical phenomena which was briefly called "world" by Minkowski is naturally four-dimensional in the space-time sense. For it is composed of individual events, each of which is described by four numbers, namely, three space co-ordinates x, y, z and a time co-ordinate, the time-value t. The "world" is in this sense also a continuum; for to every event there are as many "neighbouring" events (realised or at least thinkable) as we care to choose, the co-ordinates x_1, y_1, z_1, t_1 of which differ by an indefinitely small amount from those of the event x, y, z, t originally considered. That we have not been accustomed to regard the world in this sense as a four-dimensional continuum is due to the fact that in physics, before the advent of the theory of relativity, time played a different and more independent rôle, as compared with the space co-ordinates. It is for this reason that we have been in the habit of treating time as an independent continuum. As a matter of fact, according to classical mechanics, time is absolute, *i.e.* it is independent of the position and the condition of motion of the system of co-ordinates. We see this expressed in the last equation of the Galileian transformation ($t' = t$).

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

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

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

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

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

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

PART II

THE GENERAL THEORY OF RELATIVITY

XVIII

SPECIAL AND GENERAL PRINCIPLE OF RELATIVITY

THE basal principle, which was the pivot of all our previous considerations, was the *special* principle of relativity, *i.e.* the principle of the physical relativity of all *uniform* motion. Let us once more analyse its meaning carefully.

It was at all times clear that, from the point of view of the idea it conveys to us, every motion must only be considered as a relative motion. Returning to the illustration we have frequently used of the embankment and the railway carriage, we can express the fact of the motion here taking place in the following two forms, both of which are equally justifiable:

(*a*) The carriage is in motion relative to the embankment.
(*b*) The embankment is in motion relative to the carriage.

In (*a*) the embankment, in (*b*) the carriage, serves as the body of reference in our statement of the motion taking place. If it is simply a question of detecting or of describing the motion involved, it is in principle immaterial to what reference-body we refer the motion. As already mentioned, this is self-evident, but it must not be confused with the much more comprehensive statement called "the principle of relativity," which we have taken as the basis of our investigations.

The principle we have made use of not only maintains that we may equally well choose the carriage or the embankment as our reference-body for the description of any event (for this, too, is self-evident). Our principle rather asserts what follows: If we formulate the general laws of nature as they are obtained from experience, by making use of

(*a*) the embankment as reference-body,
(*b*) the railway carriage as reference-body,

then these general laws of nature (*e.g.* the laws of mechanics or the law of the propagation of light *in vacuo*) have exactly the same form in both cases. This can also be expressed as follows: For the *physical* description of natural processes, neither of the reference-bodies K, K' is unique (lit. "specially marked out") as compared with the other. Unlike the first, this latter statement need not of necessity hold *a priori;* it is not contained in the conceptions of "motion" and "reference-body" and derivable from them; only *experience* can decide as to its correctness or incorrectness.

Up to the present, however, we have by no means maintained the equivalence of *all* bodies of reference *K* in connection with the formulation of natural laws. Our course was more on the following lines. In the first place, we started out from the assumption that there exists a reference-body *K*, whose condition of motion is such that the Galileian law holds with respect to it: A particle left to itself and sufficiently far removed from all other particles moves uniformly in a straight line. With reference to *K* (Galileian reference-body) the laws of nature were to be as simple as possible. But in addition to *K*, all bodies of reference *K′* should be given preference in this sense, and they should be exactly equivalent to *K* for the formulation of natural laws, provided that they are in a state of *uniform rectilinear and non-rotary motion* with respect to *K*; all these bodies of reference are to be regarded as Galileian reference-bodies. The validity of the principle of relativity was assumed only for these reference-bodies, but not for others (*e.g.* those possessing motion of a different kind). In this sense we speak of the *special* principle of relativity, or special theory of relativity.

In contrast to this we wish to understand by the "general principle of relativity" the following statement: All bodies of reference K, K', etc., are equivalent for the description of natural phenomena (formulation of the general laws of nature), whatever may be their state of motion. But before proceeding farther, it ought to be pointed out that this formulation must be replaced later by a more abstract one, for reasons which will become evident at a later stage.

Since the introduction of the special principle of relativity has been justified, every intellect which strives after generalisation must feel the temptation to venture the step towards the general principle of relativity. But a simple and apparently quite reliable consideration seems to suggest that, for the present at any rate, there is little hope of success in such an attempt. Let us imagine ourselves transferred to our old friend the railway carriage, which is travelling at a uniform rate. As long as it is moving uniformly, the occupant of the carriage is not sensible of its motion, and it is for this reason that he can unreluctantly interpret the facts of the case as indicating that the carriage is at rest, but the embankment in motion. Moreover, according to the special principle of relativity, this interpretation is quite justified also from a physical point of view.

If the motion of the carriage is now changed into a non-uniform motion, as for instance by a powerful application of the brakes, then the occupant of the carriage experiences a correspondingly powerful jerk forwards. The retarded motion is manifested in the mechanical behaviour of bodies relative to the person in the railway carriage. The mechanical behaviour is different from that of the case previously considered, and for this reason it would appear to be impossible that the same mechanical laws hold relatively to the non-uniformly moving carriage, as hold with reference to the carriage when at rest or in uniform motion. At all events it is clear that the Galileian law does not hold with respect to the non-uniformly moving carriage. Because of this, we feel compelled at the present juncture to grant a kind of absolute physical reality to non-uniform motion, in opposition to the general principle of relativity. But in what follows we shall soon see that this conclusion cannot be maintained.

XIX

THE GRAVITATIONAL FIELD

"IF we pick up a stone and then let it go, why does it fall to the ground?" The usual answer to this question is: "Because it is attracted by the earth." Modern physics formulates the answer rather differently for the following reason. As a result of the more careful study of electromagnetic phenomena, we have come to regard action at a distance as a process impossible without the intervention of some intermediary medium. If, for instance, a magnet attracts a piece of iron, we cannot be content to regard this as meaning that the magnet acts directly on the iron through the intermediate empty space, but we are constrained to imagine — after the manner of Faraday — that the magnet always calls into being something physically real in the space around it, that something being what we call a "magnetic field." In its turn this magnetic field operates on the piece of iron, so that the latter strives to move towards the magnet. We shall not discuss here the justification for this incidental conception, which is indeed a somewhat arbitrary one. We shall only mention that with its aid electromagnetic phenomena can be theoretically represented much more satisfactorily than without it, and this applies particularly to the transmission of electromagnetic waves. The effects of gravitation also are regarded in an analogous manner.

The action of the earth on the stone takes place indirectly. The earth produces in its surroundings a gravitational field, which acts on the stone and produces its motion of fall. As we know from experience, the intensity of the action on a body diminishes according to a quite definite law, as we proceed farther and farther away from the earth. From our point of view this means: The law governing the properties of the gravitational field in space must be a perfectly definite one, in order correctly to represent the diminution of gravitational action with the distance from operative bodies. It is something like this: The body (*e.g.* the earth) produces a field in its immediate neighbourhood directly; the intensity and direction of the field at points farther removed from the body are thence determined by the law which governs the properties in space of the gravitational fields themselves.

In contrast to electric and magnetic fields, the gravitational field exhibits a most remarkable property, which is of fundamental importance for what follows. Bodies which are moving under the sole influence of a gravitational field receive an acceleration, *which does not in the least depend either on the material or on the physical state of the body*. For instance, a piece of lead and a piece of wood fall in exactly the same manner in a gravitational field (*in vacuo*), when they start off from rest or with the same initial velocity. This law, which holds most accurately, can be expressed in a different form in the light of the following consideration.

According to Newton's law of motion, we have

$$(\text{Force}) = (\text{inertial mass}) \times (\text{acceleration}),$$

where the "inertial mass" is a characteristic constant of the accelerated body. If now gravitation is the cause of the acceleration, we then have

$$(\text{Force}) = (\text{gravitational mass}) \times (\text{intensity of the gravitational field}),$$

where the "gravitational mass" is likewise a characteristic constant for the body. From these two relations follows:

$$(\text{acceleration}) = \frac{(\text{gravitational mass})}{(\text{inertial mass})} \times (\text{intensity of the gravitational field}).$$

If now, as we find from experience, the acceleration is to be independent of the nature and the condition of the body and always the same for a given gravitational field, then the ratio of the gravitational to the inertial mass must likewise be the same for all bodies. By a suitable choice of units we can thus make this ratio equal to unity. We then have the following law: The *gravitational* mass of a body is equal to its *inertial* mass.

It is true that this important law had hitherto been recorded in mechanics, but it had not been *interpreted*. A satisfactory interpretation can be obtained only if we recognise the following fact: *The same* quality of a body manifests itself according to circumstances as "inertia" or as "weight" (lit. "heaviness"). In the following section we shall show to what extent this is actually the case, and how this question is connected with the general postulate of relativity.

XX

THE EQUALITY OF INERTIAL AND GRAVITATIONAL MASS AS AN ARGUMENT FOR THE GENERAL POSTULATE OF RELATIVITY

WE imagine a large portion of empty space, so far removed from stars and other appreciable masses that we have before us approximately the conditions required by the fundamental law of Galilei. It is then possible to choose a Galileian reference-body for this part of space (world), relative to which points at rest remain at rest and points in motion continue permanently in uniform rectilinear motion. As reference-body let us imagine a spacious chest resembling a room with an observer inside who is equipped with apparatus. Gravitation naturally does not exist for this observer. He must fasten himself with strings to the floor, otherwise the slightest impact against the floor will cause him to rise slowly towards the ceiling of the room.

To the middle of the lid of the chest is fixed externally a hook with rope attached, and now a "being" (what kind of a being is immaterial to us) begins pulling at this with a constant force. The chest together with the observer then begin to move "upwards" with a uniformly accelerated motion. In course of time their velocity will reach unheard-of values — provided that we are viewing all this from another reference-body which is not being pulled with a rope.

But how does the man in the chest regard the process? The acceleration of the chest will be transmitted to him by the reaction of the floor of the chest. He must therefore take up this pressure by means of his legs if he does not wish to be laid out full length on the floor. He is then standing in the chest in exactly the same way as anyone stands in a room of a house on our earth. If he release a body which he previously had in his hand, the acceleration of the chest will no longer be transmitted to this body, and for this reason the body will approach the floor of the chest with an accelerated relative motion. The observer will further convince himself *that the acceleration of the body towards the floor of the chest is always of the same magnitude, whatever kind of body he may happen to use for the experiment*.

Relying on his knowledge of the gravitational field (as it was discussed in the preceding section), the man in the chest will thus come to the conclusion that he and the chest are in a gravitational field which is constant with regard to time. Of course he will be puzzled for a moment as to why the chest does not fall in this gravitational field. Just then, however, he discovers the hook in the middle of the lid of the chest and the rope which is attached to it, and he consequently comes to the conclusion that the chest is suspended at rest in the gravitational field.

Ought we to smile at the man and say that he errs in his conclusion? I do not believe we ought if we wish to remain consistent; we must rather admit that his mode of grasping the situation violates neither reason nor known mechanical laws. Even though it is being accelerated with respect to the "Galileian space" first considered, we can nevertheless regard the chest as being at rest. We have thus good grounds for extending the principle of relativity to include bodies of reference which are accelerated with respect to each other, and as a result we have gained a powerful argument for a generalised postulate of relativity.

We must note carefully that the possibility of this mode of interpretation rests on the fundamental property of the gravitational field of giving all bodies the same acceleration, or, what comes to the same thing, on the law of the equality of inertial and gravitational mass. If this natural law did not exist, the man in the accelerated chest would not be able to interpret the behaviour of the bodies around him on the supposition of a gravitational field, and he would not be justified on the grounds of experience in supposing his reference-body to be “at rest.”

Suppose that the man in the chest fixes a rope to the inner side of the lid, and that he attaches a body to the free end of the rope. The result of this will be to stretch the rope so that it will hang “vertically” downwards. If we ask for an opinion of the cause of tension in the rope, the man in the chest will say: “The suspended body experiences a downward force in the gravitational field, and this is neutralised by the tension of the rope; what determines the magnitude of the tension of the rope is the *gravitational mass* of the suspended body.” On the other hand, an observer who is poised freely in space will interpret the condition of things thus: “The rope must perforce take part in the accelerated motion of the chest, and it transmits this motion to the body attached to it. The tension of the rope is just large enough to effect the acceleration of the body. That which determines the magnitude of the tension of the rope is the *inertial mass* of the body.” Guided by this example, we see that our extension of the principle of relativity implies the *necessity* of the law of the equality of inertial and gravitational mass. Thus we have obtained a physical interpretation of this law.

From our consideration of the accelerated chest we see that a general theory of relativity must yield important results on the laws of gravitation. In point of fact, the systematic pursuit of the general idea of relativity has supplied the laws satisfied by the gravitational field. Before proceeding farther, however, I must warn the reader against a misconception suggested by these considerations. A gravitational field exists for the man in the chest, despite the fact that there was no such field for the co-ordinate system first chosen. Now we might easily suppose that the existence of a gravitational field is always only an *apparent* one. We might also think that, regardless of the kind of gravitational field which may be present, we could always choose another reference-body such that *no* gravitational field exists with reference to it. This is by no means true for all gravitational fields, but only for those of quite special form. It is, for instance, impossible to choose a body of reference such that, as judged from it, the gravitational field of the earth (in its entirety) vanishes.

We can now appreciate why that argument is not convincing, which we brought forward against the general principle of relativity at the end of Section XVIII. It is certainly true that the observer in the railway carriage experiences a jerk forwards as a result of the application of the brake, and that he recognises in this the non-uniformity of motion (retardation) of the carriage. But he is compelled by nobody to refer this jerk to a "real" acceleration (retardation) of the carriage. He might also interpret his experience thus: "My body of reference (the carriage) remains permanently at rest. With reference to it, however, there exists (during the period of application of the brakes) a gravitational field which is directed forwards and which is variable with respect to time. Under the influence of this field, the embankment together with the earth moves non-uniformly in such a manner that their original velocity in the backwards direction is continuously reduced."

XXI

IN WHAT RESPECTS ARE THE FOUNDATIONS OF CLASSICAL MECHANICS AND OF THE SPECIAL THEORY OF RELATIVITY UNSATISFACTORY?

WE have already stated several times that classical mechanics starts out from the following law: Material particles sufficiently far removed from other material particles continue to move uniformly in a straight line or continue in a state of rest. We have also repeatedly emphasised that this fundamental law can only be valid for bodies of reference *K* which possess certain unique states of motion, and which are in uniform translational motion relative to each other. Relative to other reference-bodies *K* the law is not valid. Both in classical mechanics and in the special theory of relativity we therefore differentiate between reference-bodies *K* relative to which the recognised "laws of nature" can be said to hold, and reference-bodies *K* relative to which these laws do not hold.

But no person whose mode of thought is logical can rest satisfied with this condition of things. He asks: "How does it come that certain reference-bodies (or their states of motion) are given priority over other reference-bodies (or their states of motion)? *What is the reason for this preference?* In order to show clearly what I mean by this question, I shall make use of a comparison.

I am standing in front of a gas range. Standing alongside of each other on the range are two pans so much alike that one may be mistaken for the other. Both are half full of water. I notice that steam is being emitted continuously from the one pan, but not from the other. I am surprised at this, even if I have never seen either a gas range or a pan before. But if I now notice a luminous something of bluish colour under the first pan but not under the other, I cease to be astonished, even if I have never before seen a gas flame. For I can only say that this bluish something will cause the emission of the steam, or at least *possibly* it may do so. If, however, I notice the bluish something in neither case, and if I observe that the one continuously emits steam whilst the other does not, then I shall remain astonished and dissatisfied until I have discovered some circumstance to which I can attribute the different behaviour of the two pans.

Analogously, I seek in vain for a real something in classical mechanics (or in the special theory of relativity) to which I can attribute the different behaviour of bodies considered with respect to the reference-systems K and K'.[1] Newton saw this objection and attempted to invalidate it, but without success. But E. Mach recognised it most clearly of all, and because of this objection he claimed that mechanics must be placed on a new basis. It can only be got rid of by means of a physics which is conformable to the general principle of relativity, since the equations of such a theory hold for every body of reference, whatever may be its state of motion.

[1] The objection is of importance more especially when the state of motion of the reference-body is of such a nature that it does not require any external agency for its maintenance, *e.g.* in the case when the reference-body is rotating uniformly.

XXII

A FEW INFERENCES FROM THE GENERAL THEORY* OF RELATIVITY

THE considerations of Section XX show that the general theory* of relativity puts us in a position to derive properties of the gravitational field in a purely theoretical manner. Let us suppose, for instance, that we know the space-time "course" for any natural process whatsoever, as regards the manner in which it takes place in the Galileian domain relative to a Galileian body of reference *K*. By means of purely theoretical operations (*i.e.* simply by calculation) we are then able to find how this known natural process appears, as seen from a reference-body *K′* which is accelerated relatively to *K*. But since a gravitational field exists with respect to this new body of reference *K′*, our consideration also teaches us how the gravitational field influences the process studied.

For example, we learn that a body which is in a state of uniform rectilinear motion with respect to *K* (in accordance with the law of Galilei) is executing an accelerated and in general

[* The word "theory" was changed to "principle" in both places in later editions. — J.M.]

curvilinear motion with respect to the accelerated reference-body *K'* (chest). This acceleration or curvature corresponds to the influence on the moving body of the gravitational field prevailing relatively to *K'*. It is known that a gravitational field influences the movement of bodies in this way, so that our consideration supplies us with nothing essentially new.

However, we obtain a new result of fundamental importance when we carry out the analogous consideration for a ray of light. With respect to the Galileian reference-body *K*, such a ray of light is transmitted rectilinearly with the velocity *c*. It can easily be shown that the path of the same ray of light is no longer a straight line when we consider it with reference to the accelerated chest (reference-body *K'*). From this we conclude, *that, in general, rays of light are propagated curvilinearly in gravitational fields*. In two respects this result is of great importance.

In the first place, it can be compared with the reality. Although a detailed examination of the question shows that the curvature of light rays required by the general theory of relativity is only exceedingly small for the gravitational fields at our disposal in practice, its estimated magnitude for light rays passing the sun at grazing incidence is nevertheless 1·7 seconds of arc. This ought to manifest itself in the following way.

As seen from the earth, certain fixed stars appear to be in the neighbourhood of the sun, and are thus capable of observation during a total eclipse of the sun. At such times, these stars ought to appear to be displaced outwards from the sun by an amount indicated above, as compared with their apparent position in the sky when the sun is situated at another part of the heavens. The examination of the correctness or otherwise of this deduction is a problem of the greatest importance, the early solution of which is to be expected of astronomers.[1]

In the second place our result shows that, according to the general theory of relativity, the law of the constancy of the velocity of light *in vacuo*, which constitutes one of the two fundamental assumptions in the special theory of relativity and to which we have already frequently referred, cannot claim any unlimited validity. A curvature of rays of light can only take place when the velocity of propagation of light varies with position. Now we might think that as a consequence of this, the special theory of relativity and with it the whole theory of relativity would be laid in the dust. But in reality this is not the

[1] By means of the star photographs of two expeditions equipped by a Joint Committee of the Royal and Royal Astronomical Societies, the existence of the deflection of light demanded by theory was confirmed during the solar eclipse of 29th May, 1919. (Cf. Appendix III.)

case. We can only conclude that the special theory of relativity cannot claim an unlimited domain of validity; its results hold only so long as we are able to disregard the influences of gravitational fields on the phenomena (*e.g.* of light).

Since it has often been contended by opponents of the theory of relativity that the special theory of relativity is overthrown by the general theory of relativity, it is perhaps advisable to make the facts of the case clearer by means of an appropriate comparison. Before the development of electrodynamics the laws of electrostatics and the laws of electricity were regarded indiscriminately. At the present time we know that electric fields can be derived correctly from electrostatic considerations only for the case, which is never strictly realised, in which the electrical masses are quite at rest relatively to each other, and to the co-ordinate system. Should we be justified in saying that for this reason electrostatics is overthrown by the field-equations of Maxwell in electrodynamics? Not in the least. Electrostatics is contained in electrodynamics as a limiting case; the laws of the latter lead directly to those of the former for the case in which the fields are invariable with regard to time. No fairer destiny could be allotted to any physical theory, than that it should of itself point out the way to the introduction of a more comprehensive theory, in which it lives on as a limiting case.

In the example of the transmission of light just dealt with, we have seen that the general theory of relativity enables us to derive theoretically the influence of a gravitational field on the course of natural processes, the laws of which are already known when a gravitational field is absent. But the most attractive problem, to the solution of which the general theory of relativity supplies the key, concerns the investigation of the laws satisfied by the gravitational field itself. Let us consider this for a moment.

We are acquainted with space-time domains which behave (approximately) in a "Galileian" fashion under suitable choice of reference-body, *i.e.* domains in which gravitational fields are absent. If we now refer such a domain to a reference-body K' possessing any kind of motion, then relative to K' there exists a gravitational field which is variable with respect to space and time.[1] The character of this field will of course depend on the motion chosen for K'. According to the general theory of relativity, the general law of the gravitational field must be satisfied for all gravitational fields obtainable in this way. Even though by no means all gravitational fields

[1] This follows from a generalisation of the discussion in Section XX.

can be produced in this way, yet we may entertain the hope that the general law of gravitation will be derivable from such gravitational fields of a special kind. This hope has been realised in the most beautiful manner. But between the clear vision of this goal and its actual realisation it was necessary to surmount a serious difficulty, and as this lies deep at the root of things, I dare not withhold it from the reader. We require to extend our ideas of the space-time continuum still farther.

XXIII

BEHAVIOUR OF CLOCKS AND MEASURING–RODS ON A ROTATING BODY OF REFERENCE

HITHERTO I have purposely refrained from speaking about the physical interpretation of space- and time-data in the case of the general theory of relativity. As a consequence, I am guilty of a certain slovenliness of treatment, which, as we know from the special theory of relativity, is far from being unimportant and pardonable. It is now high time that we remedy this defect; but I would mention at the outset, that this matter lays no small claims on the patience and on the power of abstraction of the reader.

We start off again from quite special cases, which we have frequently used before. Let us consider a space-time domain in which no gravitational field exists relative to a reference-body *K* whose state of motion has been suitably chosen. *K* is then a Galileian reference-body as regards the domain considered, and the results of the special theory of relativity hold relative to *K*. Let us suppose the same domain referred to a second body of reference *K′*, which is rotating uniformly with respect to *K*. In order to fix our ideas, we shall imagine *K′* to be in the form of a plane circular disc, which rotates uniformly in its own plane about its centre. An observer who is sitting eccentrically on the disc *K′* is sensible of a force which acts outwards in a radial direction, and which would be interpreted as an effect of inertia (centrifugal force) by an observer who was at rest with respect to the original reference-body *K*. But the observer on the disc may regard his disc as a reference-body which is "at rest"; on the basis of the general principle of relativity he is justified in doing this. The force acting on himself, and in fact on all other bodies which are at rest relative to the disc, he regards as the effect of a gravitational field. Nevertheless, the space-distribution of this gravitational field is of a kind that would not be possible on Newton's theory of gravitation.[1] But since the observer believes in the general theory of relativity, this does not disturb him; he is quite in the right when he believes that a general law of gravitation can be formulated — a law which not only explains the motion of the stars correctly, but also the field of force experienced by himself.

[1] The field disappears at the centre of the disc and increases proportionally to the distance from the centre as we proceed outwards.

The observer performs experiments on his circular disc with clocks and measuring-rods. In doing so, it is his intention to arrive at exact definitions for the signification of time- and space-data with reference to the circular disc *K′*, these definitions being based on his observations. What will be his experience in this enterprise?

To start with, he places one of two identically constructed clocks at the centre of the circular disc, and the other on the edge of the disc, so that they are at rest relative to it. We now ask ourselves whether both clocks go at the same rate from the standpoint of the non-rotating Galileian reference-body *K*. As judged from this body, the clock at the centre of the disc has no velocity, whereas the clock at the edge of the disc is in motion relative to *K* in consequence of the rotation. According to a result obtained in Section XII, it follows that the latter clock goes at a rate permanently slower than that of the clock at the centre of the circular disc, *i.e.* as observed from *K*. It is obvious that the same effect would be noted by an observer whom we will imagine sitting alongside his clock at the centre of the circular disc. Thus on our circular disc, or, to make the case more general, in every gravitational field, a clock will go more quickly or less quickly, according to the position in which the clock is situated (at rest). For this reason it is not possible to obtain a reasonable definition of time with the aid of clocks which are arranged at rest with respect to the body of reference. A similar difficulty presents itself when we attempt to apply our earlier definition of simultaneity in such a case, but I do not wish to go any farther into this question.

Moreover, at this stage the definition of the space co-ordinates also presents unsurmountable difficulties. If the observer applies his standard measuring-rod (a rod which is short as compared with the radius of the disc) tangentially to the edge of the disc, then, as judged from the Galileian system, the length of this rod will be less than 1, since, according to Section XII, moving bodies suffer a shortening in the direction of the motion. On the other hand, the measuring-rod will not experience a shortening in length, as judged from *K*, if it is applied to the disc in the direction of the radius. If, then, the observer first measures the circumference of the disc with his measuring-rod and then the diameter of the disc, on dividing the one by the other, he will not obtain as quotient the familiar number $\pi = 3.14 \ldots$, but a larger number,[1] whereas of course, for a disc which is at rest with respect to *K*, this operation

[1] Throughout this consideration we have to use the Galileian (non-rotating) system *K* as reference-body, since we may only assume the validity of the results of the special theory of relativity relative to *K* (relative to *K′* a gravitational field prevails).

would yield π exactly. This proves that the propositions of Euclidean geometry cannot hold exactly on the rotating disc, nor in general in a gravitational field, at least if we attribute the length 1 to the rod in all positions and in every orientation. Hence the idea of a straight line also loses its meaning. We are therefore not in a position to define exactly the co-ordinates x, y, z relative to the disc by means of the method used in discussing the special theory, and as long as the co-ordinates and times of events have not been defined we cannot assign an exact meaning to the natural laws in which these occur.

Thus all our previous conclusions based on general relativity would appear to be called in question. In reality we must make a subtle detour in order to be able to apply the postulate of general relativity exactly. I shall prepare the reader for this in the following paragraphs.

XXIV

EUCLIDEAN AND NON–EUCLIDEAN CONTINUUM

THE surface of a marble table is spread out in front of me. I can get from any one point on this table to any other point by passing continuously from one point to a "neighbouring" one, and repeating this process a (large) number of times, or, in other words, by going from point to point without executing "jumps."* I am sure the reader will appreciate with sufficient clearness what I mean here by "neighbouring" and by "jumps" (if he is not too pedantic). We express this property of the surface by describing the latter as a continuum.

Let us now imagine that a large number of little rods of equal length have been made, their lengths being small compared with the dimensions of the marble slab. When I say they are of equal length, I mean that one can be laid on any other without the ends overlapping. We next lay four of these little rods on the marble slab so that they constitute a quadrilateral figure (a square), the diagonals of which are equally long. To ensure the equality of the diagonals, we make use of a

[* jumps." — J.M.]

little testing-rod. To this square we add similar ones, each of which has one rod in common with the first. We proceed in like manner with each of these squares until finally the whole marble slab is laid out with squares. The arrangement is such, that each side of a square belongs to two squares and each corner to four squares.

It is a veritable wonder that we can carry out this business without getting into the greatest difficulties. We only need to think of the following. If at any moment three squares meet at a corner, then two sides of the fourth square are already laid, and as a consequence, the arrangement of the remaining two sides of the square is already completely determined. But I am now no longer able to adjust the quadrilateral so that its diagonals may be equal. If they are equal of their own accord, then this is an especial favour of the marble slab and of the little rods about which I can only be thankfully surprised. We must needs experience many such surprises if the construction is to be successful.

If everything has really gone smoothly, then I say that the points of the marble slab constitute a Euclidean continuum with respect to the little rod, which has been used as a "distance" (line-interval). By choosing one corner of a square as "origin," I can characterise every other corner of a square with reference to this origin by means of two numbers. I only need state how many rods I must pass over when, starting from the origin, I proceed towards the "right" and then "upwards," in order to arrive at the corner of the square under consideration. These two numbers are then the "Cartesian co-ordinates" of this corner with reference to the "Cartesian co-ordinate system" which is determined by the arrangement of little rods.

By making use of the following modification of this abstract experiment, we recognise that there must also be cases in which the experiment would be unsuccessful. We shall suppose that the rods "expand" by an amount proportional to the increase of temperature. We heat the central part of the marble slab, but not the periphery, in which case two of our little rods can still be brought into coincidence at every position on the table. But our construction of squares must necessarily come into disorder during the heating, because the little rods on the central region of the table expand, whereas those on the outer part do not.

With reference to our little rods — defined as unit lengths — the marble slab is no longer a Euclidean continuum, and we are also no longer in the position of defining Cartesian co-ordinates directly with their aid, since the above construction can no longer be carried out. But since there are other things which are not influenced in a similar manner to the little rods (or perhaps not at all) by the temperature of the table, it is possible quite naturally to maintain the point of view that the marble slab is a "Euclidean continuum." This can be done in a satisfactory manner by making a more subtle stipulation about the measurement or the comparison of lengths.

But if rods of every kind (*i.e.* of every material) were to behave *in the same way* as regards the influence of temperature when they are on the variably heated marble slab, and if we had no other means of detecting the effect of temperature than the geometrical behaviour of our rods in experiments analogous to the one described above, then our best plan would be to assign the distance *one* to two points on the slab, provided that the ends of one of our rods could be made to coincide with these two points; for how else should we define the distance without our proceeding being in the highest measure grossly arbitrary? The method of Cartesian co-ordinates must then be discarded, and replaced by another which does not assume the validity of Euclidean geometry for rigid bodies.[1] The reader will notice that

[1] Mathematicians have been confronted with our problem in the following form. If we are given a surface (*e.g.* an ellipsoid) in Euclidean three-dimensional space, then there exists for this surface a two-dimensional geometry, just as much as for a plane surface.

the situation depicted here corresponds to the one brought about by the general postulate of relativity (Section XXIII).

Gauss undertook the task of treating this two-dimensional geometry from first principles, without making use of the fact that the surface belongs to a Euclidean continuum of three dimensions. If we imagine constructions to be made with rigid rods *in the surface* (similar to that above with the marble slab), we should find that different laws hold for these from those resulting on the basis of Euclidean plane geometry. The surface is not a Euclidean continuum with respect to the rods, and we cannot define Cartesian co-ordinates *in the surface*. Gauss indicated the principles according to which we can treat the geometrical relationships in the surface, and thus pointed out the way to the method of Riemann of treating multi-dimensional, non-Euclidean *continua*. Thus it is that mathematicians long ago solved the formal problems to which we are led by the general postulate of relativity.

XXV

GAUSSIAN CO–ORDINATES

ACCORDING to Gauss, this combined analytical and geometrical mode of handling the problem can be arrived at in the following way. We imagine a system of arbitrary curves (see Fig. 4) drawn on the surface of the table. These we designate as *u*-curves, and we indicate each of them by means of a number. The curves $u=1$, $u=2$ and $u=3$ are drawn in the diagram. Between the curves $u=1$ and $u=2$ we must imagine an infinitely large number to be drawn, all of which correspond to real numbers lying between 1 and 2. We have then a system of *u*-curves, and this "infinitely dense" system covers the whole surface of the table. These *u*-curves must not intersect each other, and through each point of the surface one and only one curve must pass. Thus a perfectly definite value of *u* belongs to every point on the surface of the marble slab. In like manner we

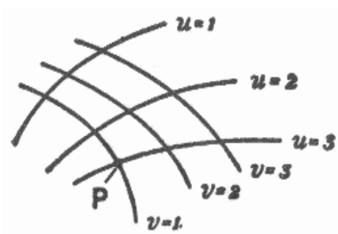


FIG. 4.

imagine a system of v-curves drawn on the surface. These satisfy the same conditions as the u-curves, they are provided with numbers in a corresponding manner, and they may likewise be of arbitrary shape. It follows that a value of u and a value of v belong to every point on the surface of the table. We call these two numbers the co-ordinates of the surface of the table (Gaussian co-ordinates). For example, the point P in the diagram has the Gaussian co-ordinates $u = 3$, $v = 1$. Two neighbouring points P and P' on the surface then correspond to the co-ordinates

$$\begin{array}{ll} P: & u, v \\ P': & u + du, v + dv, \end{array}$$

where du and dv signify very small numbers. In a similar manner we may indicate the distance (line-interval) between P and P', as measured with a little rod, by means of the very small number ds. Then according to Gauss we have

$$ds^2 = g_{11}\,du^2 + 2g_{12}\,du\,dv + g_{22}\,dv^2,$$

where g_{11}, g_{12}, g_{22}, are magnitudes which depend in a perfectly definite way on u and v. The magnitudes g_{11}, g_{12} and g_{22} determine the behaviour of the rods relative to the u-curves and v-curves, and thus also relative to the surface of the table. For the case in which the points of the surface considered form a Euclidean continuum with reference to the measuring-rods, but only in this case, it is possible to draw the u-curves and *v*-curves and to attach numbers to them, in such a manner, that we simply have:

$$ds^2 = du^2 + dv^2.$$

Under these conditions, the *u*-curves and *v*-curves are straight lines in the sense of Euclidean geometry, and they are perpendicular to each other. Here the Gaussian co-ordinates are simply Cartesian ones. It is clear that Gauss co-ordinates are nothing more than an association of two sets of numbers with the points of the surface considered, of such a nature that numerical values differing very slightly from each other are associated with neighbouring points "in space."

So far, these considerations hold for a continuum of two dimensions. But the Gaussian method can be applied also to a continuum of three, four or more dimensions. If, for instance, a continuum of four dimensions be supposed available, we may represent it in the following way. With every point of the continuum we associate arbitrarily four numbers, x_1, x_2, x_3, x_4, which are known as "co-ordinates." Adjacent points correspond to adjacent values of the co-ordinates. If a distance ds is associated with the adjacent points P and P', this distance being measurable and well-defined from a physical point of view, then the following formula holds:

$$ds^2 = g_{11}dx_1^2 + 2g_{12}dx_1\,dx_2 \;\; . \;\; . \;\; . \;\; . \;\; + g_{44}dx_4^2 \,,$$

where the magnitudes g_{11}, etc., have values which vary with the position in the continuum. Only when the continuum is a Euclidean one is it possible to associate the co-ordinates $x_1 \ldots x_4$ with the points of the continuum so that we have simply

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 .$$

In this case relations hold in the four-dimensional continuum which are analogous to those holding in our three-dimensional measurements.

However, the Gauss treatment for ds^2 which we have given above is not always possible. It is only possible when sufficiently small regions of the continuum under consideration may be regarded as Euclidean continua. For example, this obviously holds in the case of the marble slab of the table and local variation of temperature. The temperature is practically constant for a small part of the slab, and thus the geometrical behaviour of the rods is *almost* as it ought to be according to the rules of Euclidean geometry. Hence the imperfections of the construction of squares in the previous section do not show themselves clearly until this construction is extended over a considerable portion of the surface of the table.

We can sum this up as follows: Gauss invented a method for the mathematical treatment of continua in general, in which “size-relations” (“distances” between neighbouring points) are defined. To every point of a continuum are assigned as many numbers (Gaussian co-ordinates) as the continuum has dimensions. This is done in such a way, that only one meaning can be attached to the assignment, and that numbers (Gaussian co-ordinates) which differ by an indefinitely small amount are assigned to adjacent points. The Gaussian co-ordinate system is a logical generalisation of the Cartesian co-ordinate system. It is also applicable to non-Euclidean continua, but only when, with respect to the defined “size” or “distance,” small parts of the continuum under consideration behave more nearly like a Euclidean system, the smaller the part of the continuum under our notice.

XXVI

THE SPACE–TIME CONTINUUM OF THE SPECIAL THEORY OF RELATIVITY CONSIDERED AS A EUCLIDEAN CONTINUUM

WE are now in a position to formulate more exactly the idea of Minkowski, which was only vaguely indicated in Section XVII. In accordance with the special theory of relativity, certain co-ordinate systems are given preference for the description of the four-dimensional, space-time continuum. We called these “Galileian co-ordinate systems.” For these systems, the four co-ordinates x, y, z, t, which determine an event or — in other words — a point of the four-dimensional continuum, are defined physically in a simple manner, as set forth in detail in the first part of this book. For the transition from one Galileian system to another, which is moving uniformly with reference to the first, the equations of the Lorentz transformation are valid. These last form the basis for the derivation of deductions from the special theory of relativity, and in themselves they are nothing more than the expression of the universal validity of the law of transmission of light for all Galileian systems of reference.

Minkowski found that the Lorentz transformations satisfy the following simple conditions. Let us consider two neighbouring events, the relative position of which in the four-dimensional continuum is given with respect to a Galileian reference-body *K* by the space co-ordinate differences dx, dy, dz and the time-difference dt. With reference to a second Galileian system we shall suppose that the corresponding differences for these two events are dx', dy', dz', dt'. Then these magnitudes always fulfil the condition.[1]

$$dx^2 + dy^2 + dz^2 - c^2dt^2 = dx'^2 + dy'^2 + dz'^2 - c^2dt'^2.$$

The validity of the Lorentz transformation follows from this condition. We can express this as follows: The magnitude

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

which belongs to two adjacent points of the four-dimensional space-time continuum, has the same value for all selected (Galileian) reference-bodies. If we replace x, y, z, $\sqrt{-1}\,ct$, by x_1, x_2, x_3, x_4, we also obtain the result that

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2\ ^*$$

is independent of the choice of the body of refer-

[1] Cf. Appendices I and II. The relations which are derived there for the co-ordinates themselves are valid also for co-ordinate *differences*, and thus also for co-ordinate differentials (indefinitely small differences).

[* $ds^2 = dx_1^2 + dx_2^2 + dx_3^3 + dx_4^2$ — J.M.]

ence. We call the magnitude ds the "distance" apart of the two events or four-dimensional points.

Thus, if we choose as time-variable the imaginary variable $\sqrt{-1}\,ct$ instead of the real quantity t, we can regard the space-time continuum — in accordance with the special theory of relativity — as a "Euclidean" four-dimensional continuum, a result which follows from the considerations of the preceding section.

XXVII

THE SPACE–TIME CONTINUUM OF THE GENERAL THEORY OF RELATIVITY IS NOT A EUCLIDEAN CONTINUUM

IN the first part of this book we were able to make use of space-time co-ordinates which allowed of a simple and direct physical interpretation, and which, according to Section XXVI, can be regarded as four-dimensional Cartesian co-ordinates. This was possible on the basis of the law of the constancy of the velocity of light. But according to Section XXII,* the general theory of relativity cannot retain this law. On the contrary, we arrived at the result that according to this latter theory the velocity of light must always depend on the co-ordinates when a gravitational field is present. In connection with a specific illustration in Section XXIII, we found that the presence of a gravitational field invalidates the definition of the co-ordinates and the time, which led us to our objective in the special theory of relativity.

In view of the results of these considerations we are led to the conviction that, according to

[* XXI — J.M.]

the general principle of relativity, the space-time continuum cannot be regarded as a Euclidean one, but that here we have the general case, corresponding to the marble slab with local variations of temperature, and with which we made acquaintance as an example of a two-dimensional continuum. Just as it was there impossible to construct a Cartesian co-ordinate system from equal rods, so here it is impossible to build up a system (reference-body) from rigid bodies and clocks, which shall be of such a nature that measuring-rods and clocks, arranged rigidly with respect to one another, shall indicate position and time directly. Such was the essence of the difficulty with which we were confronted in Section XXIII.

But the considerations of Sections XXV and XXVI show us the way to surmount this difficulty. We refer the four-dimensional space-time continuum in an arbitrary manner to Gauss co-ordinates. We assign to every point of the continuum (event) four numbers, x_1, x_2, x_3, x_4 (co-ordinates), which have not the least direct physical significance, but only serve the purpose of numbering the points of the continuum in a definite but arbitrary manner. This arrangement does not even need to be of such a kind that we must regard x_1, x_2, x_3, as "space" co-ordinates and x_4 as a "time" co-ordinate.

The reader may think that such a description of the world would be quite inadequate. What does it mean to assign to an event the particular co-ordinates x_1, x_2, x_3, x_4, if in themselves these co-ordinates have no significance? More careful consideration shows, however, that this anxiety is unfounded. Let us consider, for instance, a material point with any kind of motion. If this point had only a momentary existence without duration, then it would be described in space-time by a single system of values x_1, x_2, x_3, x_4. Thus its permanent existence must be characterised by an infinitely large number of such systems of values, the co-ordinate values of which are so close together as to give continuity; corresponding to the material point, we thus have a (uni-dimensional) line in the four-dimensional continuum. In the same way, any such lines in our continuum correspond to many points in motion. The only statements having regard to these points which can claim a physical existence are in reality the statements about their encounters. In our mathematical treatment, such an encounter is expressed in the fact that the two lines which represent the motions of the points in question have a particular system of co-ordinate values, x_1, x_2, x_3, x_4, in common. After mature consideration the reader will doubtless admit that in reality such encounters constitute the only actual evidence of a time-space nature with which we meet in physical statements.

When we were describing the motion of a material point relative to a body of reference, we stated nothing more than the encounters of this point with particular points of the reference-body. We can also determine the corresponding values of the time by the observation of encounters of the body with clocks, in conjunction with the observation of the encounter of the hands of clocks with particular points on the dials. It is just the same in the case of space-measurements by means of measuring-rods, as a little consideration will show.

The following statements hold generally: Every physical description resolves itself into a number of statements, each of which refers to the space-time coincidence of two events *A* and *B*. In terms of Gaussian co-ordinates, every such statement is expressed by the agreement of their four co-ordinates x_1, x_2, x_3, x_4. Thus in reality, the description of the time-space continuum by means of Gauss co-ordinates completely replaces the description with the aid of a body of reference, without suffering from the defects of the latter mode of description; it is not tied down to the Euclidean character of the continuum which has to be represented.

XXVIII

EXACT FORMULATION OF THE GENERAL PRINCIPLE OF RELATIVITY

WE are now in a position to replace the provisional formulation of the general principle of relativity given in Section XVIII by an exact formulation. The form there used, "All bodies of reference K, K', etc., are equivalent for the description of natural phenomena (formulation of the general laws of nature), whatever may be their state of motion," cannot be maintained, because the use of rigid reference-bodies, in the sense of the method followed in the special theory of relativity, is in general not possible in space-time description. The Gauss co-ordinate system has to take the place of the body of reference. The following statement corresponds to the fundamental idea of the general principle of relativity: "*All Gaussian co-ordinate systems are essentially equivalent for the formulation of the general laws of nature.*"

We can state this general principle of relativity in still another form, which renders it yet more clearly intelligible than it is when in the form of the natural extension of the special principle of relativity. According to the special theory of relativity, the equations which express the general laws of nature pass over into equations of the same form when, by making use of the Lorentz transformation, we replace the space-time variables x, y, z, t, of a (Galileian) reference-body K by the space-time variables x', y', z', t', of a new reference-body K'. According to the general theory of relativity, on the other hand, by application of *arbitrary substitutions* of the Gauss variables x_1, x_2, x_3, x_4, the equations must pass over into equations of the same form; for every transformation (not only the Lorentz transformation) corresponds to the transition of one Gauss co-ordinate system into another.

If we desire to adhere to our "old-time" three-dimensional view of things, then we can characterise the development which is being undergone by the fundamental idea of the general theory of relativity as follows: The special theory of relativity has reference to Galileian domains, *i.e.* to those in which no gravitational field exists. In this connection a Galileian reference-body serves as body of reference, *i.e.* a rigid body the state of motion of which is so chosen that the Galileian law of the uniform rectilinear motion of "isolated" material points holds relatively to it.

Certain considerations suggest that we should refer the same Galileian domains to *non-Galileian* reference-bodies also. A gravitational field of a special kind is then present with respect to these bodies (cf. Sections XX and XXIII).

In gravitational fields there are no such things as rigid bodies with Euclidean properties; thus the fictitious rigid body of reference is of no avail in the general theory of relativity. The motion of clocks is also influenced by gravitational fields, and in such a way that a physical definition of time which is made directly with the aid of clocks has by no means the same degree of plausibility as in the special theory of relativity.

For this reason non-rigid reference-bodies are used which are as a whole not only moving in any way whatsoever, but which also suffer alterations in form *ad lib*. during their motion. Clocks, for which the law of motion is of any kind, however irregular, serve for the definition of time. We have to imagine each of these clocks fixed at a point on the non-rigid reference-body. These clocks satisfy only the one condition, that the "readings" which are observed simultaneously on adjacent clocks (in space) differ from each other by an indefinitely small amount. This non-rigid reference-body, which might appropriately be termed a "reference-mollusk," is in the main equivalent to a Gaussian four-dimensional co-ordinate system chosen arbitrarily. That which gives the "mollusk" a certain comprehensibleness as compared with the Gauss co-ordinate system is the (really unqualified[*]) formal retention of the separate existence of the space co-ordinates as opposed to the time co-ordinate. Every point on the mollusk is treated as a space-point, and every material point which is at rest relatively to it as at rest, so long as the mollusk is considered as reference-body. The general principle of relativity requires that all these mollusks can be used as reference-bodies with equal right and equal success in the formulation of the general laws of nature; the laws themselves must be quite independent of the choice of mollusk.

The great power possessed by the general principle of relativity lies in the comprehensive limitation which is imposed on the laws of nature in consequence of what we have seen above.

[* The word "unqualified" was correctly changed to "unjustified" in later editions. — J.M.]

XXIX

THE SOLUTION OF THE PROBLEM OF GRAVITATION ON THE BASIS OF THE GENERAL PRINCIPLE OF RELATIVITY

IF the reader has followed all our previous considerations, he will have no further difficulty in understanding the methods leading to the solution of the problem of gravitation.

We start off from a consideration of a Galileian domain, *i.e.* a domain in which there is no gravitational field relative to the Galileian reference-body *K*. The behaviour of measuring-rods and clocks with reference to *K* is known from the special theory of relativity, likewise the behaviour of "isolated" material points; the latter move uniformly and in straight lines.

Now let us refer this domain to a random Gauss co-ordinate system or to a "mollusk" as reference-body *K′*. Then with respect to *K′* there is a gravitational field *G* (of a particular kind). We learn the behaviour of measuring-rods and clocks and also of freely-moving material points with reference to *K′* simply by mathematical transformation. We interpret this behaviour as the behaviour of measuring-rods, clocks and material points under the influence of the gravitational field *G*. Hereupon we introduce a hypothesis: that the influence of the gravitational field on measuring-rods, clocks and freely-moving material points continues to take place according to the same laws, even in the case when the prevailing gravitational field is *not* derivable from the Galileian special case, simply by means of a transformation of co-ordinates.

The next step is to investigate the space-time behaviour of the gravitational field *G*, which was derived from the Galileian special case simply by transformation of the co-ordinates. This behaviour is formulated in a law, which is always valid, no matter how the reference-body (mollusk) used in the description may be chosen.

This law is not yet the *general* law of the gravitational field, since the gravitational field under consideration is of a special kind. In order to find out the general law-of-field of gravitation we still require to obtain a generalisation of the law as found above. This can be obtained without caprice, however, by taking into consideration the following demands:

(*a*) The required generalisation must likewise satisfy the general postulate of relativity.

(*b*) If there is any matter in the domain under consideration, only its inertial mass, and thus according to Section XV only its energy is of importance for its effect in exciting a field.

(*c*) Gravitational field and matter together must satisfy the law of the conservation of energy (and of impulse).

Finally, the general principle of relativity permits us to determine the influence of the gravitational field on the course of all those processes which take place according to known laws when a gravitational field is absent, *i.e.* which have already been fitted into the frame of the special theory of relativity. In this connection we proceed in principle according to the method which has already been explained for measuring-rods, clocks and freely-moving material points.

The theory of gravitation derived in this way from the general postulate of relativity excels not only in its beauty; nor in removing the defect attaching to classical mechanics which was brought to light in Section XXI; nor in interpreting the empirical law of the equality of inertial and gravitational mass; but it has also already explained a result of observation in astronomy, against which classical mechanics is powerless.

If we confine the application of the theory to the case where the gravitational fields can be regarded as being weak, and in which all masses move with respect to the co-ordinate system with velocities which are small compared with the velocity of light, we then obtain as a first approximation the Newtonian theory. Thus the latter theory is obtained here without any particular assumption, whereas Newton had to introduce the hypothesis that the force of attraction between mutually attracting material points is inversely proportional to the square of the distance between them. If we increase the accuracy of the calculation, deviations from the theory of Newton make their appearance, practically all of which must nevertheless escape the test of observation owing to their smallness.

We must draw attention here to one of these deviations. According to Newton's theory, a planet moves round the sun in an ellipse, which would permanently maintain its position with respect to the fixed stars, if we could disregard the motion of the fixed stars themselves and the action of the other planets under consideration. Thus, if we correct the observed motion of the planets for these two influences, and if Newton's theory be strictly correct, we ought to obtain for the orbit of the planet an ellipse, which is fixed with reference to the fixed stars. This deduction, which can be tested with great accuracy, has been confirmed for all the planets save one, with the precision that is capable of being obtained by the delicacy of observation attainable at the present time. The sole exception is Mercury, the planet which lies nearest the sun. Since the time of Leverrier, it has been known that the ellipse corresponding to the orbit of Mercury, after it has been corrected for the influences mentioned above, is not stationary with respect to the fixed stars, but that it rotates exceedingly slowly in the plane of the orbit and in the sense of the orbital motion. The value obtained for this rotary movement of the orbital ellipse was 43 seconds of arc per century, an amount ensured to be correct to within a few seconds of arc. This effect can be explained by means of classical mechanics only on the assumption of hypotheses which have little probability, and which were devised solely for this purpose.

On the basis of the general theory of relativity, it is found that the ellipse of every planet round the sun must necessarily rotate in the manner indicated above; that for all the planets, with the exception of Mercury, this rotation is too small to be detected with the delicacy of observation possible at the present time; but that in the case of Mercury it must amount to 43 seconds of arc per century, a result which is strictly in agreement with observation.

Apart from this one, it has hitherto been possible to make only two deductions from the theory which admit of being tested by observation, to wit, the curvature of light rays by the gravitational field of the sun,[1] and a displacement of the spectral lines of light reaching us from large stars, as compared with the corresponding lines for light produced in an analogous manner terrestrially (*i.e.* by the same kind of molecule[*]). I do not doubt that these deductions from the theory will be confirmed also.

[1] Observed by Eddington and others in 1919. (Cf. Appendix III.)

[[*] The word “molecule” was correctly changed to “atom” in later editions. Cf. Appendix III, pg. 157. — J.M.]

PART III

CONSIDERATIONS ON THE UNIVERSE AS A WHOLE

XXX

COSMOLOGICAL DIFFICULTIES OF NEWTON'S THEORY

APART from the difficulty discussed in Section XXI, there is a second fundamental difficulty attending classical celestial mechanics, which, to the best of my knowledge, was first discussed in detail by the astronomer Seeliger. If we ponder over the question as to how the universe, considered as a whole, is to be regarded, the first answer that suggests itself to us is surely this: As regards space (and time) the universe is infinite. There are stars everywhere, so that the density of matter, although very variable in detail, is nevertheless on the average everywhere the same. In other words: However far we might travel through space, we should find everywhere an attenuated swarm of fixed stars of approximately the same kind and density.

This view is not in harmony with the theory of Newton. The latter theory rather requires that the universe should have a kind of centre in which the density of the stars is a maximum, and that as we proceed outwards from this centre the group-density of the stars should diminish, until finally, at great distances, it is succeeded by an infinite region of emptiness. The stellar universe ought to be a finite island in the infinite ocean of space.[1]

This conception is in itself not very satisfactory. It is still less satisfactory because it leads to the result that the light emitted by the stars and also individual stars of the stellar system are perpetually passing out into infinite space, never to return, and without ever again coming into interaction with other objects of nature. Such a finite material universe would be destined to become gradually but systematically impoverished.

[1] *Proof.* — According to the theory of Newton, the number of "lines of force" which come from infinity and terminate in a mass m is proportional to the mass m. If, on the average, the mass-density ρ_0 is constant throughout the universe, then a sphere of volume V will enclose the average mass $\rho_0 V$. Thus the number of lines of force passing through the surface F of the sphere into its interior is proportional to $\rho_0 V$. For unit area of the surface of the sphere the number of lines of force which enters the sphere is thus proportional to $\rho_0 \frac{V}{F}$ * or $\rho_0 R$. Hence the intensity of the field at the surface would ultimately become infinite with increasing radius R of the sphere, which is impossible.

[* $\rho_0 \frac{v}{F}$ — J.M.]

In order to escape this dilemma, Seeliger suggested a modification of Newton's law, in which he assumes that for great distances the force of attraction between two masses diminishes more rapidly than would result from the inverse square law. In this way it is possible for the mean density of matter to be constant everywhere, even to infinity, without infinitely large gravitational fields being produced. We thus free ourselves from the distasteful conception that the material universe ought to possess something of the nature of a centre. Of course we purchase our emancipation from the fundamental difficulties mentioned, at the cost of a modification and complication of Newton's law which has neither empirical nor theoretical foundation. We can imagine innumerable laws which would serve the same purpose, without our being able to state a reason why one of them is to be preferred to the others; for any one of these laws would be founded just as little on more general theoretical principles as is the law of Newton.

XXXI

THE POSSIBILITY OF A "FINITE" AND YET "UNBOUNDED" UNIVERSE

BUT speculations on the structure of the universe also move in quite another direction. The development of non-Euclidean geometry led to the recognition of the fact, that we can cast doubt on the *infiniteness* of our space without coming into conflict with the laws of thought or with experience (Riemann, Helmholtz). These questions have already been treated in detail and with unsurpassable lucidity by Helmholtz and Poincaré, whereas I can only touch on them briefly here.

In the first place, we imagine an existence in two-dimensional space. Flat beings with flat implements, and in particular flat rigid measuring-rods, are free to move in a *plane*. For them nothing exists outside of this plane: that which they observe to happen to themselves and to their flat "things" is the all-inclusive reality of their plane. In particular, the constructions of plane Euclidean geometry can be carried out by means of the rods, *e.g.* the lattice construction, considered in Section XXIV. In contrast to ours, the universe of these beings is two-dimensional; but, like ours, it extends to infinity. In their universe there is room for an infinite number of identical squares made up of rods, *i.e.* its volume (surface) is infinite. If these beings say their universe is "plane," there is sense in the statement, because they mean that they can perform the constructions of plane Euclidean geometry with their rods. In this connection the individual rods always represent the same distance, independently of their position.

Let us consider now a second two-dimensional existence, but this time on a spherical surface instead of on a plane. The flat beings with their measuring-rods and other objects fit exactly on this surface and they are unable to leave it. Their whole universe of observation extends exclusively over the surface of the sphere. Are these beings able to regard the geometry of their universe as being plane geometry and their rods withal as the realisation of "distance"? They cannot do this. For if they attempt to realise a straight line, they will obtain a curve, which we "three-dimensional beings" designate as a great circle, *i.e.* a self-contained line of definite finite length, which can be measured up by means of a measuring-rod. Similarly, this universe has a finite area, that can be compared with the area of a square constructed with rods. The great charm resulting from this consideration lies in the recognition of the fact that *the universe of these beings is finite and yet has no limits*.

But the spherical-surface beings do not need to go on a world-tour in order to perceive that they are not living in a Euclidean universe. They can convince themselves of this on every part of their "world," provided they do not use too small a piece of it. Starting from a point, they draw "straight lines" (arcs of circles as judged in three-dimensional space) of equal length in all directions. They will call the line joining the free ends of these lines a "circle." For a plane surface, the ratio of the circumference of a circle to its diameter, both lengths being measured with the same rod, is, according to Euclidean geometry of the plane, equal to a constant value π, which is independent of the diameter of the circle. On their spherical surface our flat beings would find for this ratio the value

$$\pi\frac{\sin\left(\frac{r}{R}\right)}{\left(\frac{r}{R}\right)},$$

i.e. a smaller value than π, the difference being the more considerable, the greater is the radius of the circle in comparison with the radius R of the "world-sphere." By means of this relation the spherical beings can determine the radius of their universe ("world"), even when only a relatively small part of their world-sphere is available for their measurements. But if this part is very small indeed, they will no longer be able to demonstrate that they are on a spherical "world" and not on a Euclidean plane, for a small part of a spherical surface differs only slightly from a piece of a plane of the same size.

Thus if the spherical-surface beings are living on a planet of which the solar system occupies only a negligibly small part of the spherical universe, they have no means of determining whether they are living in a finite or in an infinite universe, because the "piece of universe" to which they have access is in both cases practically plane, or Euclidean. It follows directly from this discussion, that for our sphere-beings the circumference of a circle first increases with the radius until the "circumference of the universe" is reached, and that it thenceforward gradually decreases to zero for still further increasing values of the radius. During this process the area of the circle continues to increase more and more, until finally it becomes equal to the total area of the whole "world-sphere."

Perhaps the reader will wonder why we have placed our "beings" on a sphere rather than on another closed surface. But this choice has its justification in the fact that, of all closed surfaces, the sphere is unique in possessing the property that all points on it are equivalent. I admit that the ratio of the circumference *c* of a circle to its radius *r* depends on *r*, but for a given value of *r* it is the same for all points of the "world-sphere"; in other words, the "world-sphere" is a "surface of constant curvature."

To this two-dimensional sphere-universe there is a three-dimensional analogy, namely, the three-dimensional spherical space which was discovered by Riemann. Its points are likewise all equivalent. It possesses a finite volume, which is determined by its "radius" ($2\pi^2 R^3$). Is it possible to imagine a spherical space? To imagine a space means nothing else than that we imagine an epitome of our "space" experience, *i.e.* of experience that we can have in the movement of "rigid" bodies. In this sense we *can* imagine a spherical space.

Suppose we draw lines or stretch strings in all directions from a point, and mark off from each of these the distance *r* with a measuring-rod. All the free end-points of these lengths lie on a spherical surface. We can specially measure up the area (F) of this surface by means of a square made up of measuring-rods. If the universe is Euclidean, then $F = 4\pi r^2$; if it is spherical, then F is always less than $4\pi r^2$. With increasing values of r, F increases from zero up to a maximum value which is determined by the "world-radius," but for still further increasing values of r, the area gradually diminishes to zero. At first, the straight lines which radiate from the starting point diverge farther and farther from one another, but later they approach each other, and finally they run together again at a "counter-point" to the starting point. Under such conditions they have traversed the whole spherical space. It is easily seen that the three-dimensional spherical space is quite analogous to the two-dimensional spherical surface. It is finite (*i.e.* of finite volume), and has no bounds.

It may be mentioned that there is yet another kind of curved space: "elliptical space." It can be regarded as a curved space in which the two "counter-points" are identical (indistinguishable from each other). An elliptical universe can thus be considered to some extent as a curved universe possessing central symmetry.

It follows from what has been said, that closed spaces without limits are conceivable. From amongst these, the spherical space (and the elliptical) excels in its simplicity, since all points on it are equivalent. As a result of this discussion, a most interesting question arises for astronomers and physicists, and that is whether the universe in which we live is infinite, or whether it is finite in the manner of the spherical universe. Our experience is far from being sufficient to enable us to answer this question. But the general theory of relativity permits of our answering it with a moderate degree of certainty, and in this connection the difficulty mentioned in Section XXX finds its solution.

XXXII

THE STRUCTURE OF SPACE ACCORDING TO THE GENERAL THEORY OF RELATIVITY

ACCORDING to the general theory of relativity, the geometrical properties of space are not independent, but they are determined by matter. Thus we can draw conclusions about the geometrical structure of the universe only if we base our considerations on the state of the matter as being something that is known. We know from experience that, for a suitably chosen co-ordinate system, the velocities of the stars are small as compared with the velocity of transmission of light. We can thus as a rough approximation arrive at a conclusion as to the nature of the universe as a whole, if we treat the matter as being at rest.

We already know from our previous discussion that the behaviour of measuring-rods and clocks is influenced by gravitational fields, *i.e.* by the distribution of matter. This in itself is sufficient to exclude the possibility of the exact validity of Euclidean geometry in our universe. But it is conceivable that our universe differs only slightly from a Euclidean one, and this notion seems all the more probable, since calculations show that the metrics of surrounding space is influenced only to an exceedingly small extent by masses even of the magnitude of our sun. We might imagine that, as regards geometry, our universe behaves analogously to a surface which is irregularly curved in its individual parts, but which nowhere departs appreciably from a plane: something like the rippled surface of a lake. Such a universe might fittingly be called a quasi-Euclidean universe. As regards its space it would be infinite. But calculation shows that in a quasi-Euclidean universe the average density of matter would necessarily be *nil*. Thus such a universe could not be inhabited by matter everywhere; it would present to us that unsatisfactory picture which we portrayed in Section XXX.

If we are to have in the universe an average density of matter which differs from zero, however small may be that difference, then the universe cannot be quasi-Euclidean. On the contrary, the results of calculation indicate that if matter be distributed uniformly, the universe would necessarily be spherical (or elliptical). Since in reality the detailed distribution of matter is not uniform, the real universe will deviate in individual parts from the spherical, *i.e.* the universe will be quasi-spherical. But it will be necessarily finite. In fact, the theory supplies us with a simple connection [1] between the space-expanse of the universe and the average density of matter in it.

[1] For the "radius" R of the universe we obtain the equation

$$R^2 = \frac{2}{\kappa\rho}$$

The use of the C.G.S. system in this equation gives $\frac{2}{\kappa} = 1{\cdot}08 . 10^{27}$; ρ is the average density of the matter.

APPENDIX I

SIMPLE DERIVATION OF THE LORENTZ TRANSFORMATION [SUPPLEMENTARY TO SECTION XI]

FOR the relative orientation of the co-ordinate systems indicated in Fig. 2, the x-axes of both systems permanently coincide. In the present case we can divide the problem into parts by considering first only events which are localised on the x-axis. Any such event is represented with respect to the co-ordinate system K by the abscissa x and the time t, and with respect to the system K' by the abscissa x' and the time t'. We require to find x' and t' when x and t are given.

A light-signal, which is proceeding along the positive axis of x, is transmitted according to the equation

$$x = ct$$

or

$$x - ct = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1).$$

Since the same light-signal has to be transmitted relative to K' with the velocity c, the propagation relative to the system K' will be represented by the analogous formula

$$x' - ct' = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2).$$

Those space-time points (events) which satisfy (1) must also satisfy (2). Obviously this will be the case when the relation

$$(x' - ct') = \lambda(x - ct) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

is fulfilled in general, where λ indicates a constant; for, according to (3), the disappearance of $(x - ct)$ involves the disappearance of $(x' - ct')$.

If we apply quite similar considerations to light rays which are being transmitted along the negative x-axis, we obtain the condition

$$(x' + ct') = \mu(x + ct) \quad . \quad . \quad . \quad . \quad . \quad . \quad (4).$$

By adding (or subtracting) equations (3) and (4), and introducing for convenience the constants a and b in place of the constants λ and μ where

$$a = \frac{\lambda + \mu}{2}$$

and

$$b = \frac{\lambda - \mu}{2},$$

we obtain the equations

$$\left.\begin{aligned} x' &= ax - bct \\ ct' &= act - bx \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5).$$

We should thus have the solution of our problem, if the constants a and b were known. These result from the following discussion.

For the origin of K' we have permanently $x' = 0$, and hence according to the first of the equations (5)

$$x = \frac{bc}{a} t.$$

If we call v the velocity with which the origin of K' is moving relative to K, we then have

$$v = \frac{bc}{a} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6).$$

The same value v can be obtained from equations[*] (5), if we calculate the velocity of another point of K' relative to K, or the velocity (directed towards the negative x-axis) of a point of K with respect to K'. In short, we can designate v as the relative velocity of the two systems.

Furthermore, the principle of relativity teaches us that, as judged from K, the length of a unit measuring-rod which is at rest with reference to K' must be exactly the same as the length, as judged from K', of a unit measuring-rod which is at rest relative to K. In order to see how the points of the x'-axis appear as viewed from K, we only require to take a "snapshot" of K' from K; this means that we have to insert a particular value of t (time of K), *e.g.*[†] $t = 0$. For this value of t we then obtain from the first of the equations (5)

$$x' = ax.$$

Two points of the x'-axis which are separated by the distance $\Delta x' = 1$[‡] when measured in the K' system are thus separated in our instantaneous photograph by the distance

$$\Delta x = \frac{1}{a} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7).$$

[* tion — J.M.] [† e.g. — J.M.] [‡ $x' = 1$ — J.M.]

But if the snapshot be taken from K' ($t' = 0$), and if we eliminate t from the equations (5), taking into account the expression (6), we obtain

$$x' = a\left(1 - \frac{v^2}{c^2}\right)x.$$

From this we conclude that two points on the x-axis and separated by the distance 1 (relative to K) will be represented on our snapshot by the distance

$$\Delta x' = a\left(1 - \frac{v^2}{c^2}\right) \quad . \; . \; . \; . \; . \; . \quad (7a).$$

But from what has been said, the two snapshots must be identical; hence Δx in (7) must be equal to $\Delta x'$ in ($7a$), so that we obtain

$$a^2 = \frac{1}{1 - \frac{v^2}{c^2}} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (7b).$$

The equations (6) and ($7b$) determine the constants a and b. By inserting the values of these constants in (5), we obtain the first and the fourth of the equations given in Section XI.

$$\left.\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}} \\ t' &= \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned}\right\} \quad . \; . \; . \; . \; . \; . \; . \quad (8).$$

Thus we have obtained the Lorentz transformation for events on the x-axis. It satisfies the condition

$$x'^2 - c^2t'^2 = x^2 - c^2t^2 \quad \ldots \ldots \quad (8a).$$

The extension of this result, to include events which take place outside the x-axis, is obtained by retaining equations (8) and supplementing them by the relations

$$\left.\begin{aligned} y' &= y \\ z' &= z \end{aligned}\right\} \quad \ldots \ldots \ldots \quad (9).$$

In this way we satisfy the postulate of the constancy of the velocity of light *in vacuo* for rays of light of arbitrary direction, both for the system K and for the system K'. This may be shown in the following manner.

We suppose a light-signal sent out from the origin of K at the time $t = 0$. It will be propagated according to the equation

$$r = \sqrt{x^2 + y^2 + z^2} = ct,$$

or, if we square this equation, according to the equation

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \quad \ldots \ldots \quad (10).$$

It is required by the law of propagation of light, in conjunction with the postulate of relativity, that the transmission of the signal in question should take place — as judged from K' — in accordance with the corresponding formula

$$r' = ct'$$

or,

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \quad \ldots \ldots \quad (10a).$$

In order that equation (10*a*) may be a consequence of equation (10), we must have

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = \sigma(x^2 + y^2 + z^2 - c^2t^2) \quad (11).$$

Since equation (8*a*) must hold for points on the *x*-axis, we thus have $\sigma = 1$. It is easily seen that the Lorentz transformation really satisfies equation (11) for $\sigma = 1$; for (11) is a consequence of (8*a*) and (9), and hence also of (8) and (9). We have thus derived the Lorentz transformation.

The Lorentz transformation represented by (8) and (9) still requires to be generalised. Obviously it is immaterial whether the axes of *K'* be chosen so that they are spatially parallel to those of *K*. It is also not essential that the velocity of translation of *K'* with respect to *K* should be in the direction of the *x*-axis. A simple consideration shows that we are able to construct the Lorentz transformation in this general sense from two kinds of transformations, viz. from Lorentz transformations in the special sense and from purely spatial transformations, which corresponds to the replacement of the rectangular co-ordinate system by a new system with its axes pointing in other directions.

Mathematically, we can characterise the generalised Lorentz transformation thus:

It expresses x', y', z', t', in terms of linear homogeneous functions of x, y, z, t, of such a kind that the relation

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2 \ . \quad (11a)$$

is satisfied identically. That is to say: If we substitute their expressions in x, y, z, t, in place of x', y', z', t', on the left-hand side, then the left-hand side of (11a) agrees with the right-hand side.

APPENDIX II

MINKOWSKI'S FOUR — DIMENSIONAL SPACE ("WORLD") [Supplementary to Section XVII]

WE can characterise the Lorentz transformation still more simply if we introduce the imaginary $\sqrt{-1} \cdot ct$ in place of t, as time-variable. If, in accordance with this, we insert

$$\begin{aligned} x_1 &= x \\ x_2 &= y \\ x_3 &= z \\ x_4 &= \sqrt{-1} \cdot ct, \end{aligned}$$

and similarly for the accented system K', then the condition which is identically satisfied by the transformation can be expressed thus:

$$x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2. \quad (12).$$

That is, by the afore-mentioned choice of "co-ordinates" (11*a*) is transformed into this equation.

We see from (12) that the imaginary time co-ordinate x_4 enters into the condition of transformation in exactly the same way as the space co-ordinates x_1, x_2, x_3. It is due to this fact that, according to the theory of relativity, the "time"

x_4 enters into natural laws in the same form as the space co-ordinates x_1, x_2, x_3.

A four-dimensional continuum described by the "co-ordinates" x_1, x_2, x_3, x_4, was called "world" by Minkowski, who also termed a point-event a "world-point." From a "happening" in three-dimensional space, physics becomes, as it were, an "existence" in the four-dimensional "world."

This four-dimensional "world" bears a close similarity to the three-dimensional "space" of (Euclidean) analytical geometry. If we introduce into the latter a new Cartesian co-ordinate system (x'_1, x'_2, x'_3) with the same origin, then x'_1, x'_2, x'_3, are linear homogeneous functions of x_1, x_2, x_3, which identically satisfy the equation

$$x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + x_3^2.$$

The analogy with (12) is a complete one. We can regard Minkowski's "world" in a formal manner as a four-dimensional Euclidean space (with imaginary time co-ordinate); the Lorentz transformation corresponds to a "rotation" of the co-ordinate system in the four-dimensional "world."

APPENDIX III

THE EXPERIMENTAL CONFIRMATION OF THE GENERAL THEORY OF RELATIVITY

FROM a systematic theoretical point of view, we may imagine the process of evolution of an empirical science to be a continuous process of induction. Theories are evolved, and are expressed in short compass as statements of a large number of individual observations in the form of empirical laws, from which the general laws can be ascertained by comparison. Regarded in this way, the development of a science bears some resemblance to the compilation of a classified catalogue. It is, as it were, a purely empirical enterprise.

But this point of view by no means embraces the whole of the actual process; for it slurs over the important part played by intuition and deductive thought in the development of an exact science. As soon as a science has emerged from its initial stages, theoretical advances are no longer achieved merely by a process of arrangement. Guided by empirical data, the investigator rather develops a system of thought which, in general, is built up logically from a small number of fundamental assumptions, the so-called axioms. We call such a system of thought a *theory*. The theory finds the justification for its existence in the fact that it correlates a large number of single observations, and it is just here that the "truth" of the theory lies.

Corresponding to the same complex of empirical data, there may be several theories, which differ from one another to a considerable extent. But as regards the deductions from the theories which are capable of being tested, the agreement between the theories may be so complete, that it becomes difficult to find such deductions in which the two theories differ from each other. As an example, a case of general interest is available in the province of biology, in the Darwinian theory of the development of species by selection in the struggle for existence, and in the theory of development which is based on the hypothesis of the hereditary transmission of acquired characters.

We have another instance of far-reaching agreement between the deductions from two theories in Newtonian mechanics on the one hand, and the general theory of relativity on the other. This agreement goes so far, that up to the present we have been able to find only a few deductions from the general theory of relativity which are capable of investigation, and to which the physics of pre-relativity days does not also lead, and this despite the profound difference in the fundamental assumptions of the two theories. In what follows, we shall again consider these important deductions, and we shall also discuss the empirical evidence appertaining to them which has hitherto been obtained.

(*a*) Motion of the Perihelion of Mercury

According to Newtonian mechanics and Newton's law of gravitation, a planet which is revolving round the sun would describe an ellipse round the latter, or, more correctly, round the common centre of gravity of the sun and the planet. In such a system, the sun, or the common centre of gravity, lies in one of the foci of the orbital ellipse in such a manner that, in the course of a planet-year, the distance sun-planet grows from a minimum to a maximum, and then decreases again to a minimum. If instead of Newton's law we insert a somewhat different law of attraction into the calculation, we find that, according to this new law, the motion would still take place in such a manner that the distance sun-planet exhibits periodic variations; but in this case the angle described by the line joining sun and planet during such a period (from perihelion — closest proximity to the sun — to perihelion) would differ from 360°. The line of the orbit would not then be a closed one, but in the course of time it would fill up an annular part of the orbital plane, viz. between the circle of least and the circle of greatest distance of the planet from the sun.

According also to the general theory of relativity, which differs of course from the theory of Newton, a small variation from the Newton-Kepler motion of a planet in its orbit should take place, and in such a way, that the angle described by the radius sun-planet between one perihelion and the next should exceed that corresponding to one complete revolution by an amount given by

$$+\frac{24\pi^3 a^2}{T^2 c^2 (1-e^2)}.$$

(*N.B.* — One complete revolution corresponds to the angle 2π in the absolute angular measure customary in physics, and the above expression gives the amount by which the radius sun-planet exceeds this angle during the interval between one perihelion and the next.) In this expression a represents the major semi-axis of the ellipse, e its eccentricity, c the velocity of light, and T the period of revolution of the planet. Our result may also be stated as follows: According to the general theory of relativity, the major axis of the ellipse rotates round the sun in the same sense as the orbital motion of the planet. Theory requires that this rotation should amount to 43 seconds of arc per century for the planet Mercury, but for the other planets of our solar system its magnitude should be so small that it would necessarily escape detection.[1]

In point of fact, astronomers have found that the theory of Newton does not suffice to calculate the observed motion of Mercury with an exactness corresponding to that of the delicacy of observation attainable at the present time. After taking account of all the disturbing influences exerted on Mercury by the remaining planets, it was found (Leverrier — 1859 — and Newcomb — 1895) that an unexplained perihelial movement of the orbit of Mercury remained over, the amount of which does not differ sensibly from the above-mentioned + 43 seconds of arc per century. The uncertainty of the empirical result amounts to a few seconds only.

(*b*) Deflection of Light by a Gravitational Field

In Section XXII it has been already mentioned that, according to the general theory of relativity, a ray of light will experience a curvature of its

[1] Especially since the next planet Venus has an orbit that is almost an exact circle, which makes it more difficult to locate the perihelion with precision.

path when passing through a gravitational field, this curvature being similar to that experienced by the path of a body which is projected through a gravitational field. As a result of this theory, we should expect that a ray of light which is passing close to a heavenly body would be deviated towards the latter. For a ray of light which passes the sun at a distance of Δ sun-radii from its centre, the angle of deflection (α) should amount to

$$\alpha = \frac{1 \cdot 7 \text{ seconds of arc}}{\Delta}.$$

It may be added that, according to the theory, half of this deflection is produced by the Newtonian field of attraction of the sun, and the other half by the geometrical modification ("curvature") of space caused by the sun.

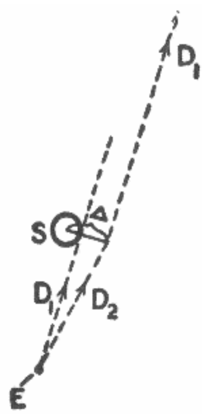


FIG. 5.

This result admits of an experimental test by means of the photographic registration of stars during a total eclipse of the sun. The only reason why we must wait for a total eclipse is because at every other time the atmosphere is so strongly illuminated by the light from the sun that the stars situated near the sun's disc are invisible. The predicted effect can be seen clearly from the accompanying diagram. If the sun (S) were not present, a star which is practically infinitely distant would be seen in the direction D_1, as observed from the earth. But as a consequence of the deflection of light from the star by the sun, the star will be seen in the direction D_2, *i.e.* at a somewhat greater distance from the centre of the sun than corresponds to its real position.

In practice, the question is tested in the following way. The stars in the neighbourhood of the sun are photographed during a solar eclipse.

In addition, a second photograph of the same stars is taken when the sun is situated at another position in the sky, *i.e.* a few months earlier or later. As compared with the standard photograph, the positions of the stars on the eclipse-photograph ought to appear displaced radially outwards (away from the centre of the sun) by an amount corresponding to the angle α.

We are indebted to the Royal Society and to the Royal Astronomical Society for the investigation of this important deduction. Undaunted by the war and by difficulties of both a material and a psychological nature aroused by the war, these societies equipped two expeditions — to Sobral (Brazil) and to the island of Principe (West Africa) — and sent several of Britain's most celebrated astronomers (Eddington, Cottingham, Crommelin, Davidson), in order to obtain photographs of the solar eclipse of 29th May, 1919. The relative discrepancies to be expected between the stellar photographs obtained during the eclipse and the comparison photographs amounted to a few hundredths of a millimetre only. Thus great accuracy was necessary in making the adjustments required for the taking of the photographs, and in their subsequent measurement.

The results of the measurements confirmed the theory in a thoroughly satisfactory manner. The rectangular components of the observed and of the calculated deviations of the stars (in seconds of arc) are set forth in the following table of results:

Number of the Star.	First Co-ordinate.		Second Co-ordinate.	
	Observed.	Calculated.	Observed.	Calculated.
11 . .	– 0·19	– 0·22	+ 0·16	+ 0·02
5 . .	+ 0·29	+ 0·31	– 0·46	– 0·43
4 . .	+ 0·11	+ 0·10	+ 0·83	+ 0·74
3 . .	+ 0·20	+ 0·12	+ 1·00	+ 0·87
6 . .	+ 0·10	+ 0·04	+ 0·57	+ 0·40
10 . .	– 0·08	+ 0·09	+ 0·35	+ 0·32
2 . .	+ 0·95	+ 0·85	– 0·27	– 0·09

(*c*) Displacement of Spectral Lines towards the Red

In Section XXIII it has been shown that in a system K' which is in rotation with regard to a Galileian system K, clocks of identical construction, and which are considered at rest with respect to the rotating reference-body, go at rates which are dependent on the positions of the clocks. We shall now examine this dependence quantitatively. A clock, which is situated at a distance *r* from the centre of the disc, has a velocity relative to *K* which is given by

$$v = \omega r,$$

where ω represents the[*] velocity of rotation of the disc *K′* with respect to *K*. If ν_0 represents the number of ticks of the clock per unit time ("rate" of the clock) relative to *K* when the clock is at rest, then the "rate" of the clock (ν) when it is moving relative to *K* with a velocity *v*, but at rest with respect to the disc, will, in accordance with Section XII, be given by

$$\nu = \nu_0 \sqrt{1 - \frac{v^2}{c^2}},$$

or with sufficient accuracy by

$$\nu = \nu_0 \left(1 - \tfrac{1}{2}\frac{v^2}{c^2}\right).$$

This expression may also be stated in the following form:

$$\nu = \nu_0 \left(1 - \frac{1}{c^2}\frac{\omega^2 r^2}{2}\right).$$

If we represent the difference of potential of the centrifugal force between the position of the clock and the centre of the disc by ϕ, *i.e.* the work,

[* The word "angular" was inserted here in later editions. — J.M.]

considered negatively, which must be performed on the unit of mass against the centrifugal force in order to transport it from the position of the clock on the rotating disc to the centre of the disc, then we have

$$\phi = -\frac{\omega^2 r^2}{2}.$$

From this it follows that

$$\nu = \nu_0 \left(1 + \frac{\phi}{c^2}\right).$$

In the first place, we see from this expression that two clocks of identical construction will go at different rates when situated at different distances from the centre of the disc. This result is also valid from the standpoint of an observer who is rotating with the disc.

Now, as judged from the disc, the latter is in a gravitational field of potential ϕ, hence the result we have obtained will hold quite generally for gravitational fields. Furthermore, we can regard an atom which is emitting spectral lines as a clock, so that the following statement will hold:

An atom absorbs or emits light of a frequency which is dependent on the potential of the gravitational field in which it is situated.

The frequency of an atom situated on the surface of a heavenly body will be somewhat less than the frequency of an atom of the same element which is situated in free space (or on the surface of a smaller celestial body).

Now $\phi = -K\frac{M}{r}$, where K is Newton's constant of gravitation, and M is the mass of the heavenly body. Thus a displacement towards the red ought to take place for spectral lines produced at the surface of stars as compared with the spectral lines of the same element produced at the surface of the earth, the amount of this displacement being

$$\frac{\nu_0 - \nu}{\nu_0} = \frac{K}{c^2}\frac{M}{r}.$$

For the sun, the displacement towards the red predicted by theory amounts to about two millionths of the wave-length. A trustworthy calculation is not possible in the case of the stars, because in general neither the mass M nor the radius r is known.

It is an open question whether or not this effect exists, and at the present time astronomers are working with great zeal towards the solution. Owing to the smallness of the effect in the case of the sun, it is difficult to form an opinion as to its existence. Whereas Grebe and Bachem (Bonn), as a result of their own measurements and those of Evershed and Schwarzschild on the cyanogen bands, have placed the existence of the effect almost beyond doubt, other investigators, particularly St. John, have been led to the opposite opinion in consequence of their measurements.

Mean displacements of lines towards the less refrangible end of the spectrum are certainly revealed by statistical investigations of the fixed stars; but up to the present the examination of the available data does not allow of any definite decision being arrived at, as to whether or not these displacements are to be referred in reality to the effect of gravitation. The results of observation have been collected together, and discussed in detail from the standpoint of the question which has been engaging our attention here, in a paper by E. Freundlich entitled "Zur Prüfung der allgemeinen Relativitäts-Theorie" (*Die Naturwissenschaften*, 1919, No. 35, p. 520: Julius Springer, Berlin).

At all events, a definite decision will be reached during the next few years. If the displacement of spectral lines towards the red by the gravitational potential does not exist, then the general theory of relativity will be untenable. On the other hand, if the cause of the displacement of spectral lines be definitely traced to the gravitational potential, then the study of this displacement will furnish us with important information as to the mass of the heavenly bodies.

BIBLIOGRAPHY

WORKS IN ENGLISH ON EINSTEIN'S THEORY

INTRODUCTORY

The Foundations of Einstein's Theory of Gravitation: Erwin Freundlich (translation by H. L. Brose). Camb. Univ. Press, 1920.

Space and Time in Contemporary Physics: Moritz Schlick (translation by H. L. Brose). Clarendon Press, Oxford, 1920.

THE SPECIAL THEORY

The Principle of Relativity: E. Cunningham. Camb. Univ. Press.

Relativity and the Electron Theory: E. Cunningham, Monographs on Physics. Longmans, Green & Co.

The Theory of Relativity: L. Silberstein. Macmillan & Co.

The Space-Time Manifold of Relativity: E. B. Wilson and G. N. Lewis, *Proc. Amer. Soc. Arts & Science*, vol. xlviii., No. 11, 1912.

THE GENERAL THEORY

Report on the Relativity Theory of Gravitation: A. S. Eddington. Fleetway Press Ltd., Fleet Street, London.

On Einstein's Theory of Gravitation and its Astronomical Consequences: W. de Sitter, *M. N. Roy. Astron. Soc.*, lxxvi. p. 699, 1916; lxxvii. p. 155, 1916; lxxviii. p. 3, 1917.

On Einstein's Theory of Gravitation: H. A. Lorentz, *Proc. Amsterdam Acad.*, vol. xix. p. 1341, 1917.

Space, Time and Gravitation: W. de Sitter: *The Observatory*, No. 505, p. 412. Taylor & Francis, Fleet Street, London.

The Total Eclipse of 29th May 1919, and the Influence of Gravitation on Light: A. S. Eddington, *ibid.*, March, 1919.

Discussion on the Theory of Relativity: *M. N. Roy. Astron. Soc.*, vol. lxxx., No. 2., p. 96, December 1919.

The Displacement of Spectrum Lines and the Equivalence Hypothesis: W. G. Duffield, *M. N. Roy. Astron. Soc.*, vol. lxxx.; No. 3, p. 262, 1920.

Space, Time and Gravitation: A. S. Eddington. Camb. Univ. Press, 1920.

ALSO, CHAPTERS IN

The Mathematical Theory of Electricity and Magnetism: J. H. Jeans (4th edition). Camb. Univ. Press, 1920.

The Electron Theory of Matter: O. W. Richardson. Camb. Univ. Press.

INDEX

INDEX

[* 170 — J.M.]

''JUST THE MATHS''

by

A.J. Hobson

TEACHING UNITS - TABLE OF CONTENTS

(Average number of pages = 1038 ¸140 = 7.4 per unit)

All units are in presented as .PDF files

(Home) (Foreword) (About the Author)

Page last changed: 3 October 2002
Contact for this page: C J Judd

FOREWORD

(Home) (About the Author) (Teaching Units) (Teaching Slides)

In 35 years of teaching mathematics to Engineers and Scientists, I have frequently been made aware (by students) of a common cry for help. "We're coping, generally, with our courses", they may say, "but it's Just the Maths". This is the title chosen for the package herein.

Traditional text-books and programmed learning texts can sometimes include a large amount of material which is not always needed for a particular course; and which can leave students feeling that there is too much to cope with. Many such texts are biased towards the mathematics required for specific engineering or scientific disciplines and emphasise the associated practical applications in their lists of tutorial examples. There can also be a higher degree of mathematical rigor than would be required by students who are not intending to follow a career in mathematics itself.

"Just the Maths" is a collection of separate units, in chronological topic-order, intended to service foundation level and first year degree level courses in higher education, especially those delivered in a modular style. Each unit represents, on average, the work to be covered in a typical two-hour session consisting of a lecture and a tutorial. However, since each unit attempts to deal with self-contained and, where possible, independent topics, it may sometimes require either more than or less than two hours spent on it.

"Just the Maths" does not have the format of a traditional text-book or a course of programmed learning; but it is written in a traditional pure-mathematics style with the minimum amount of formal rigor. By making use of the well-worn phrase, "it can be shown that", it is able to concentrate on the core mathematical techniques required by any scientist or engineer. The techniques are demonstrated by worked examples and reinforced by exercises that are few enough in number to allow completion, or near-completion, in a one-hour tutorial session. Answers to exercises are supplied at the end of each unit of work.

A.J. Hobson
January 2002

Page last changed: 9 September 2002
Contact for this page: C J Judd

ABOUT THE AUTHOR

Tony Hobson was, until retirement in November 2001, a Senior Lecturer in Mathematics of the School of Mathematical and Information Sciences at Coventry University. He graduated from the University College of Wales, Aberystwyth in 1964, with a BSc. Degree 2(i) in Pure Mathematics, and from Birmingham University in 1965, with an MSc. Degree in Pure Mathematics. His Dissertation for the MSc. Degree consisted of an investigation into the newer styles teaching Mathematics in the secondary schools of the 1960's with the advent of experiments such as the Midland Mathematics Experiment and the School Mathematics Project. His teaching career began in 1965 at the Rugby College of Engineering Technology where, as well as involvement with the teaching of Analysis and Projective Geometry to the London External Degree in Mathematics, he soon developed a particular interest in the teaching of Mathematics to Science and Engineering Students. This interest continued after the creation of the Polytechnics in 1971 and a subsequent move to the Coventry Polytechnic, later to become Coventry University. It was his main teaching interest throughout the thirty six years of his career; and it meant that much of the time he spent on research and personal development was in the area of curriculum development. In 1982 he became a Non-stipendiary Priest in the Church of England, an interest he maintained throughout his retirement. Tony Hobson died in December 2002.

The set of teaching units for "Just the Maths" has been the result of a pruning, honing and computer-processing exercise (over some four or five years) of **many** years' personal teaching materials, into a form which may be easily accessible to students of Science and Engineerng in the future.

Page last changed: 17 January 2003
Contact for this page: C J Judd

"JUST THE MATHS"

UNIT NUMBER

1.1

ALGEBRA 1
(Introduction to algebra)

by

A.J. Hobson

UNIT 1.1 - ALGEBRA 1 - INTRODUCTION TO ALGEBRA

DEFINITION
An **"Algebra"** is any Mathematical system which uses the concepts of Equality (=), Addition (+), Subtraction (-), Multiplication (× or .) and Division (÷).

Note:
The Algebra of Numbers is what we normally call **"Arithmetic"** and, as far as this unit is concerned, it is only the algebra of numbers which we shall be concerned with.

1.1.1 THE LANGUAGE OF ALGEBRA

Suppose we use the symbols a, b and c to denote numbers of arithmetic; then

(a) $a + b$ is called the **"sum of a and b"**.

Note:
$a + a$ is usually abbreviated to $2a$,
$a + a + a$ is usually abbreviated to $3a$ and so on.

(b) $a - b$ is called the **"difference between a and b"**.

(c) $a \times b$, $a.b$ or even just ab is called the **"product"** of a and b.

Notes:
(i)
$a.a$ is usually abbreviated to a^2,
$a.a.a$ is usually abbreviated to a^3 and so on.

(ii) $-1 \times a$ is usually abbreviated to $-a$ and is called the **"negation"** of a.

(d) $a \div b$ or $\frac{a}{b}$ is called the **"quotient"** or **"ratio"** of a and b.

(e) $\frac{1}{a}$, [also written a^{-1}], is called the **"reciprocal"** of a.

(f) $\mid a \mid$ is called the **"modulus"**, **"absolute value"** or **"numerical value"** of a. It can be defined by the two statements

$\mid a \mid = a$ when a is positive or zero;
$\mid a \mid = -a$ when a is negative or zero.

Note:
Further work on fractions (ratios) will appear later, but we state here for reference the rules for combining fractions together:

Rules for combining fractions together

1.
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

2.
$$\frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd}$$

3.
$$\frac{a}{b}.\frac{c}{d} = \frac{a.c}{b.d}$$

4.
$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \times \frac{d}{c} = \frac{a.d}{b.c}$$

EXAMPLES

1. How much more than the difference of 127 and 59 is the sum of 127 and 59 ?

   Solution
   The difference of 127 and 59 is $127-59 = 68$ and the sum of 127 and 59 is $127+59 = 186$. The sum exceeds the difference by $186 - 68 = 118$.

2. What is the reciprocal of the number which is 5 multiplied by the difference of 8 and 2 ?

   Solution
   We require the reciprocal of $5.(8 - 2)$; that is, the reciprocal of 30. The answer is therefore $\frac{1}{30}$.

3. Calculate the value of $4\frac{2}{3} - 5\frac{1}{9}$ expressing the answer as a fraction.

   Solution
   Converting both numbers to a single fraction, we require
   $$\frac{14}{3} - \frac{46}{9} = \frac{126 - 138}{27} = -\frac{12}{27} = -\frac{4}{9}.$$
   We could also have observed that the 'lowest common multiple' (see later) of the two denominators, 3 and 9, is 9; hence we could write the alternative solution
   $$\frac{42}{9} - \frac{46}{9} = -\frac{4}{9}.$$

4. Remove the modulus signs from the expression $\mid a - 2 \mid$ in the cases when (i) a is greater than (or equal to) 2 and (ii) a is less than 2.

   Solution
   (i) If a is greater than or equal to 2,

$$\mid a - 2 \mid = a - 2;$$

   (ii) If a is less than 2,

$$\mid a - 2 \mid = -(a - 2) = 2 - a.$$

1.1.2 THE LAWS OF ALGEBRA

If the symbols a, b and c denote numbers of arithmetic, then the following Laws are obeyed by them:

(a) The Commutative Law of Addition $\quad a + b = b + a$

(b) The Associative Law of Addition $\quad a + (b + c) = (a + b) + c$

(c) The Commutative Law of Multiplication $\quad a.b = b.a$

(d) The Associative Law of Multipication $\quad a.(b.c) = (a.b).c$

(e) The Distributive Laws $\quad a.(b + c) = a.b + a.c$ and $(a + b).c = a.c + b.c$

Notes:

(i) A consequence of the Distributive Laws is the rule for multiplying together a pair of bracketted expressions. It will be encountered more formally later, but we state it here for reference:

$$(a + b).(c + d) = a.c + b.c + a.d + b.d$$

(ii) The alphabetical letters so far used for numbers in arithmetic have been taken from the **beginning** of the alphabet. These tend to be reserved for fixed quantities called **constants**. Letters from the **end** of the alphabet, such as w, x, y, z are normally used for quantities which may take many values, and are called **variables**.

1.1.3 PRIORITIES IN CALCULATIONS

Suppose that we encountered the expression $5 \times 6 - 4$. It would seem to be ambiguous, meaning either $30 - 4 = 26$ or $5 \times 2 = 10$.

However, we may remove the ambiguity by using brackets where necessary, together with a rule for precedence between the use of the brackets and the symbols $+$, $-$, $\times$ and $\div$.

The rule is summarised in the abbreviation

B.O.D.M.A.S.

which means that the order of precedence is

B	brackets	( )	First Priority
O	of	$\times$	Joint Second Priority
D	division	$\div$	Joint Second Priority
M	multiplication	$\times$	Joint Second Priority
A	addition	$+$	Joint Third Priority
S	subtraction	$-$	Joint Third Priority

Thus, $5 \times (6 - 4) = 5 \times 2 = 10$
but $5 \times 6 - 4 = 30 - 4 = 26$.

Similarly, $12 \div 3 - 1 = 4 - 1 = 3$
whereas $12 \div (3 - 1) = 12 \div 2 = 6$.

1.1.4 FACTORS

If a number can be expressed as a product of other numbers, each of those other numbers is called a **"factor"** of the original number.

EXAMPLES

1. We may observe that
$$70 = 2 \times 7 \times 5$$
so that the number 70 has factors of 2, 7 and 5. These three cannot be broken down into factors themselves because they are what are known as **"prime"** numbers (numbers whose only factors are themselves and 1). Hence the only factors of 70, apart from 70 and 1, are 2, 7 and 5.

2. Show that the numbers 78 and 182 have two common factors which are prime numbers.

   The two factorisations are as follows:
$$78 = 2 \times 3 \times 13,$$

$$182 = 2 \times 7 \times 13.$$

The common factors are thus 2 and 13, both of which are prime numbers.

Notes:

(i) If two or more numbers have been expressed as a product of their prime factors, we may easily identify the prime factors which are common to all the numbers and hence obtain the **"highest common factor"**, h.c.f.

For example, $90 = 2 \times 3 \times 3 \times 5$ and $108 = 2 \times 2 \times 3 \times 3 \times 3$. Hence the h.c.f $= 2 \times 3 \times 3 = 18$

(ii) If two or more numbers have been expressed as a product of their prime factors, we may also identify the **"lowest common multiple"**, l.c.m.

For example, $15 = 3 \times 5$ and $20 = 2 \times 2 \times 5$. Hence the smallest number into which both 15 and 20 will divide requires two factors of 2 (for 20), one factor of 5 (for both 15 and 20) and one factor of 3 (for 15). The l.c.m. is thus $2 \times 2 \times 3 \times 5 = 60$.

(iii) If the numerator and denominator of a fraction have factors in common, then such factors may be cancelled to leave the fraction in its **"lowest terms"**.

For example $\frac{15}{105} = \frac{3\times5}{3\times5\times7} = \frac{1}{7}$.

1.1.5 EXERCISES

1. Find the sum and product of
   (a) 3 and 6; (b) 10 and 7; (c) 2, 3 and 6;
   (d) $\frac{3}{2}$ and $\frac{4}{11}$; (e) $1\frac{2}{5}$ and $7\frac{3}{4}$; (f) $2\frac{1}{7}$ and $5\frac{4}{21}$.
2. Find the difference between and quotient of
   (a) 18 and 9; (b) 20 and 5; (c) 100 and 20;
   (d) $\frac{3}{5}$ and $\frac{7}{10}$; (e) $3\frac{1}{4}$ and $2\frac{2}{9}$; (f) $1\frac{2}{3}$ and $5\frac{5}{6}$.
3. Evaluate the following expressions:
   (a) $6 - 2 \times 2$; (b) $(6 - 2) \times 2$;
   (c) $6 \div 2 - 2$; (d) $(6 \div 2) - 2$;
   (e) $6 - 2 + 3 \times 2$; (f) $6 - (2 + 3) \times 2$;
   (g) $(6 - 2) + 3 \times 2$; (h) $\frac{16}{-2}$; (i) $\frac{-24}{-3}$; (j) $(-6) \times (-2)$.

4. Place brackets in the following to make them correct:
   (a) $6 \times 12 - 3 + 1 = 55$; (b) $6 \times 12 - 3 + 1 = 68$;
   (c) $6 \times 12 - 3 + 1 = 60$; (d) $5 \times 4 - 3 + 2 = 7$;
   (e) $5 \times 4 - 3 + 2 = 15$; (f) $5 \times 4 - 3 + 2 = -5$.

5. Express the following as a product of prime factors:

   (a) 26; (b) 100; (c) 27; (d) 71;

   (e) 64; (f) 87; (g) 437; (h) 899.

6. Find the h.c.f of

   (a) 12, 15 and 21; (b) 16, 24 and 40; (c) 28, 70, 120 and 160;

   (d) 35, 38 and 42; (e) 96, 120 and 144.

7. Find the l.c.m of

   (a) 5, 6, and 8; (b) 20 and 30; (c) 7, 9 and 12;

   (d) 100, 150 and 235; (e) 96, 120 and 144.

1.1.6 ANSWERS TO EXERCISES

1. (a) 9, 18; (b) 17,70; (c) 11,36; (d) $\frac{41}{22}$, $\frac{6}{11}$; (e) $\frac{183}{20}$, $\frac{217}{20}$; (f) $\frac{154}{21}$, $\frac{545}{49}$.

2. (a) 9,2; (b) 15,4; (c) 80,5; (d) $-\frac{1}{10}$, $\frac{6}{7}$; (e) $\frac{37}{36}$, $\frac{117}{80}$; (f) $-\frac{25}{6}$, $\frac{2}{7}$.

3. (a) 2; (b) 8; (c) 1; (d) 1; (e) 10;
   (f) -4; (g) 10; (h) -8; (i) 8; (j) 12;

4. (a) $6 \times (12 - 3) + 1 = 55$; (b) $6 \times 12 - (3 + 1) = 68$;
   (c) $6 \times (12 - 3 + 1) = 60$; (d) $5 \times (4 - 3) + 2 = 7$;
   (e) $5 \times 4 - (3 + 2) = 15$; (f) $5 \times (4 - [3 + 2]) = -5$.

5. (a) 2×13; (b) $2 \times 2 \times 5 \times 5$; (c) $3 \times 3 \times 3$; (d) 71×1;

   (e) $2 \times 2 \times 2 \times 2 \times 2 \times 2$; (f) 3×29; (g) 19×23; (h) 29×31.

6. (a) 3; (b) 8; (c) 2; (d) 1; (e) 24.

7. (a) 120; (b) 60; (c) 252; (d) 14100; (e) 1440.

"JUST THE MATHS"

UNIT NUMBER

1.2

ALGEBRA 2
(Numberwork)

by

A.J. Hobson

UNIT 1.2 - - ALGEBRA 2 - NUMBERWORK

1.2.1 TYPES OF NUMBER

In this section (and elsewhere) the meaning of the following types of numerical quantity will need to be appreciated:

(a) NATURAL NUMBERS
These are the counting numbers 1, 2, 3, 4, ...

(b) INTEGERS
These are the positive and negative whole numbers and zero;
i.e. ......$-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5,$ ......

(c) RATIONALS
These are the numbers which can be expressed as the ratio of two integers but can also be written as a terminating or recurring decimal (see also next section)

For example

$$\frac{2}{5} = 0.4$$

and

$$\frac{3}{7} = 0.428714287142871....$$

(d) IRRATIONALS
These are the numbers which cannot be expressed as either the ratio of two integers or a recurring decimal (see also next section)

Typical examples are numbers like

$$\pi \simeq 3.1415926.....$$

$$e \simeq 2.71828.....$$

$$\sqrt{2} \simeq 1.4142135....$$

$$\sqrt{5} \simeq 2.2360679....$$

The above four types of number form the system of **"real numbers"**.

1.2.2 DECIMAL NUMBERS

(a) Rounding to a specified number of decimal places
Most decimal quantities used in scientific work need to be approximated by **"rounding"** them (up or down as appropriate) to a specified number of decimal places, depending on the accuracy required.

When rounding to n decimal places, the digit in the n-th place is left as it is when the one after it is below 5; otherwise it is taken up by one digit.

EXAMPLES

1. 362.5863 = 362.586 to 3 decimal places;
   362.5863 = 362.59 to 2 decimal places;
   362.5863 = 362.6 to 1 decimal place;
   362.5863 = 363 to the nearest whole number.
2. 0.02158 = 0.0216 to 4 decimal places;
   0.02158 = 0.022 to three decimal places;
   0.02158 = 0.02 to 2 decimal places.

(b) Rounding to a specified number of significant figures
The first significant figure of a decimal quantity is the first non-zero digit from the left, whether it be before or after the decimal point.

Hence when rounding to a specified number of significant figures, we use the same principle as in **(a)**, but starting from the first significant figure, then working to the right.

EXAMPLES

1. 362.5863 = 362.59 to 5 significant figures;
   362.5863 = 362.6 to 4 significant figures;
   362.5863 = 363 to 3 significant figures;
   362.5863 = 360 to 2 significant figures;
   362.5863 = 400 to 1 significant figure.
2. 0.02158 = 0.0216 to 3 significant figures; 0.02158 = 0.022 to 2 significant figures;
   0.02158 = 0.02 to 1 significant figure.

1.2.3 THE USE OF ELECTRONIC CALCULATORS

(a) B.O.D.M.A.S.
The student will normally need to work to the instruction manual for the particular calculator being used; but care must be taken to remember the B.O.D.M.A.S. rule for priorities in calculations when pressing the appropriate buttons.

For example, in working out $7.25 + 3.75 \times 8.32$, the multiplication should be carried out first, then the addition. The answer is 38.45, not 91.52.

Similarly, in working out $6.95 \div [2.43 - 1.62]$, it is best to evaluate $2.43 - 1.62$, then generate its reciprocal with the $\frac{1}{x}$ button, then multiply by 6.95. The answer is 8.58, not 1.24

(b) Other Useful Numerical Functions
Other useful functions to become familiar with for scientific work with numbers are those indicated by labels such as $\sqrt{x}$, x^2, x^y and $x^{\frac{1}{y}}$, using, where necessary, the **"shift"** control to bring the correct function into operation.

For example:

$$\sqrt{173} \simeq 13.153;$$
$$173^2 = 29929;$$
$$23^3 = 12167;$$
$$23^{\frac{1}{3}} \simeq 2.844$$

(c) The Calculator Memory
Familiarity with the calculator's memory facility will be essential for more complicated calculations in which various parts need to be stored temporarily while the different steps are being carried out.

For example, in order to evaluate

$$(1.4)^3 - 2(1.4)^2 + 5(1.4) - 3 \simeq 2.824$$

we need to store each of the four terms in the calculation (positively or negatively) then recall their total sum at the end.

1.2.4 SCIENTIFIC NOTATION

(a) Very large numbers, especially decimal numbers are customarily written in the form

$$a \times 10^n$$

where n is a positive integer and a lies between 1 and 10.

For instance,

$$521983677.103 = 5.21983677103 \times 10^8.$$

(b) Very small decimal numbers are customarily written in the form

$$a \times 10^{-n}$$

where n is a positive integer and a lies between 1 and 10.

For instance,

$$0.00045938 = 4.5938 \times 10^{-4}.$$

Note:
An electronic calculator will allow you to enter numbers in scientific notation by using the $\boxed{\textbf{EXP}}$ or $\boxed{\textbf{EE}}$ buttons.

EXAMPLES

1. Key in the number 3.90816×10^{57} on a calculator.

   Press **3.90816** $\boxed{\textbf{EXP}}$ **57**

   In the display there will now be 3.90816 57 or $3.90816^{\times 10^{57}}$.

2. Key in the number 1.5×10^{-27} on a calculator

   Press **1.5** $\boxed{\textbf{EXP}}$ **27** $\boxed{+/-}$

   In the display there will now be 1.5 - 27 or $1.5^{\times 10^{-27}}$.

Notes:

(i) On a calculator or computer, scientific notation is also called *floating point notation.*

(ii) When performing a calculation involving decimal numbers, it is always a good idea to check that the result is reasonable and that a major arithmetical error has not been made with the calculator.

For example,

$$69.845 \times 196.574 = 6.9845 \times 10^1 \times 1.96574 \times 10^3.$$

This product can be **estimated** for reasonableness as:

$$7 \times 2 \times 1000 = 14000.$$

The answer obtained by calculator is 13729.71 to two decimal places which is 14000 when rounded to the nearest 1000, indicating that the exact result could be reasonably expected.

(iii) If a set of measurements is made with an accuracy to a given number of significant figures, then it may be shown that any calculation involving those measurements will be accurate only to one significant figure more than the least number of significant figures in any measurement.

For example, the edges of a rectangular piece of cardboard are measured as 12.5cm and 33.43cm respectively and hence the area may be evaluated as

$$12.5 \times 33.43 = 417.875\text{cm}^2.$$

Since one of the edges is measured only to three significant figures, the area result is accurate only to four significant figures and hence must be stated as 417.9cm^2.

1.2.5 PERCENTAGES

Definition
A percentage is a fraction whose denominator is 100. We use the per-cent symbol, %, to represent a percentage.

For instance, the fraction $\frac{17}{100}$ may be written 17%

EXAMPLES

1. Express $\frac{2}{5}$ as a percentage.
   Solution
$$\frac{2}{5} = \frac{2}{5} \times \frac{20}{20} = \frac{40}{100} = 40\%$$

2. Calculate 27% of 90.
   Solution
$$27\% \text{ of } 90 = \frac{27}{100} \times 90 = \frac{27}{10} \times 9 = 24.3$$

3. Express 30% as a decimal.

   Solution

$$30\% = \frac{30}{100} = 0.3$$

1.2.6 RATIO

Sometimes, a more convenient way of expressing the ratio of two numbers is to use a colon (:) in place of either the standard division sign (÷) or the standard notation for fractions.

For instance, the expression 7:3 could be used instead of either $7 \div 3$ or $\frac{7}{3}$. It denotes that two quantities are "in the ratio 7 to 3" which implies that the first number is seven thirds times the second number or, alternatively, the second number is three sevenths times the first number. Although more cumbersome, the ratio 7:3 could also be written $\frac{7}{3}$:1 or 1:$\frac{3}{7}$.

EXAMPLES

1. Divide 170 in the ratio 3:2

   Solution

   We may consider that 170 is made up of $3 + 2 = 5$ parts, each of value $\frac{170}{5} = 34$.

   Three of these make up a value of $3 \times 34 = 102$ and two of them make up a value of $2 \times 34 = 68$.

   Thus 170 needs to be divided into 102 and 68.

2. Divide 250 in the ratio 1:3:4

   Solution

   This time, we consider that 250 is made up of $1 + 3 + 4 = 8$ parts, each of value $\frac{250}{8} = 31.25$. Three of these make up a value of $3 \times 31.25 = 93.75$ and four of them make up a value of $4 \times 31.25 = 125$.

   Thus 250 needs to be divided into 31.25, 93.75 and 125.

1.2.7 EXERCISES

1. Write to 3 s.f.

   (a) 6962; (b) 70.406; (c) 0.0123;
   (d) 0.010991; (e) 45.607; (f) 2345.

2. Write 65.999 to

   (a) 4 s.f. (b) 3 s.f. (c) 2 s.f.
   (d) 1 s.f. (e) 2 d.p. (f) 1 d.p.

3. Compute the following in scientific notation:

   (a) $(0.003)^2 \times (0.00004) \times (0.00006) \times 5,000,000,000$;

   (b) $800 \times (0.00001)^2 \div (200,000)^4$.

4. Assuming that the following contain numbers obtained by measurement, use a calculator to determine their value and state the expected level of accuracy:

   (a)
   $$\frac{(13.261)^{0.5}(1.2)}{(5.632)^3};$$

   (b)
   $$\frac{(8.342)(-9.456)^3}{(3.25)^4}.$$

5. Calculate 23% of 124.

6. Express the following as percentages:

   (a) $\frac{9}{11}$; (b) $\frac{15}{20}$; (c) $\frac{9}{10}$; (d) $\frac{45}{50}$; (e) $\frac{75}{90}$.

7. A worker earns £400 a week, then receives a 6% increase. Calculate the new weekly wage.

8. Express the following percentages as decimals:

   (a) 50% (b) 36% (c) 75% (d) 100% (e) 12.5%

9. Divide 180 in the ratio 8:1:3

10. Divide 930 in the ratio 1:1:3

11. Divide 6 in the ratio 2:3:4

12. Divide 1200 in the ratio 1:2:3:4

13. A sum of £2600 is to be divided in the ratio $2\frac{3}{4} : 1\frac{1}{2} : 2\frac{1}{4}$. Calculate the amount of money in each part of the division.

1.2.8. ANSWERS TO EXERCISES

1. (a) 6960; (b) 70.4; (c) 0.0123;
   (d) 0.0110; (e) 45.6; (f) 2350.
2. (a) 66.00; (b) 66.0; (c) 66;
   (d) 70; (e) 66.00; (f) 66.0
3. (a) 1.08 - 04 or $1.08^{\times 10^{-4}}$ (b) 5 - 29 or $5^{\times 10^{-29}}$;
4. (a) 0.0245, accurate to three sig. figs. (b) -63.22, accurate to four sig. figs.
5. 28.52
6. (a) 81.82% (b) 75% (c) 90% (d) 90% (e) 83.33%
7. £424.
8. (a) 0.5; (b) 0.36; (c) 0.75; (d) 1; (e) 0.125
9. 120, 15, 45.
10. 186, 186, 558.
11. 1.33, 2, 2.67
12. 120, 240, 360, 480
13. £1100, £600 £900.

"JUST THE MATHS"

UNIT NUMBER

1.3

ALGEBRA 3
(Indices and radicals (or surds))

by

A.J.Hobson

UNIT 1.3 - ALGEBRA 3 - INDICES AND RADICALS (or Surds)

1.3.1 INDICES

(a) Positive Integer Indices

It was seen earlier that, for any number a, a^2 denotes $a.a$, a^3 denotes a.a.a, a^4 denotes $a.a.a.a$ and so on.

Suppose now that a and b are arbitrary numbers and that m and n are natural numbers (i.e. positive whole numbers)

Then the following rules are the basic Laws of Indices:

Law No. 1

$$a^m \times a^n = a^{m+n}$$

Law No. 2

$$a^m \div a^n = a^{m-n}$$

assuming, for the moment, that m is greater than n.

Note:
It is natural to use this rule to give a definition to a^0 which would otherwise be meaningless.

Clearly $\frac{a^m}{a^m} = 1$ but the present rule for indices suggests that $\frac{a^m}{a^m} = a^{m-m} = a^0$.
Hence, we **define** a^0 to be equal to 1.

Law No. 3

$$(a^m)^n = a^{mn}$$

$$a^m b^m = (ab)^m$$

EXAMPLE

Simplify the expression,

$$\frac{x^2y^3}{z} \div \frac{xy}{z^5}.$$

Solution

The expression becomes

$$\frac{x^2y^3}{z} \times \frac{z^5}{xy} = xy^2z^4.$$

(b) Negative Integer Indices

Law No. 4

$$a^{-1} = \frac{1}{a}$$

Note:
It has already been mentioned that a^{-1} means the same as $\frac{1}{a}$; and the logic behind this statement is to maintain the basic Laws of Indices for negative indices as well as positive ones.

For example $\frac{a^m}{a^{m+1}}$ is clearly the same as $\frac{1}{a}$ but, using Law No. 2 above, it could also be thought of as $a^{m-[m+1]} = a^{-1}$.

Law No. 5

$$a^{-n} = \frac{1}{a^n}$$

Note:
This time, we may observe that $\frac{a^m}{a^{m+n}}$ is clearly the same as $\frac{1}{a^n}$; but we could also use Law No. 2 to interpret it as $a^{m-[m+n]} = a^{-n}$

Law No. 6

$$a^{-\infty} = 0$$

Note:
Strictly speaking, no power of a number can ever be equal to zero, but Law No. 6 asserts that a very large negative power of a number a gives a very small value; the larger the negative power, the smaller will be the value.

EXAMPLE

Simplify the expression,

$$\frac{x^5y^2z^{-3}}{x^{-1}y^4z^5} \div \frac{z^2x^2}{y^{-1}}.$$

Solution

The expression becomes

$$x^5y^2z^{-3}xy^{-4}z^{-5}y^{-1}z^{-2}x^{-2} = x^4y^{-3}z^{-10}.$$

(c) Rational Indices

(i) Indices of the form $\frac{1}{n}$ where n is a natural number.

In order to preserve Law No. 3, we interpret $a^{\frac{1}{n}}$ to mean a number which gives the value a when it is raised to the power n. It is called an **"n-th Root of a"** and, sometimes there is more than one value.

ILLUSTRATION

$$81^{\frac{1}{4}} = \pm 3 \text{ but } (-27)^{\frac{1}{3}} = -3 \text{ only.}$$

(ii) Indices of the form $\frac{m}{n}$ where m and n are natural numbers with no common factor.

The expression $y^{\frac{m}{n}}$ may be interpreted in two ways as either $(y^m)^{\frac{1}{n}}$ or $(y^{\frac{1}{n}})^m$. It may be shown that both interpretations give the same result but, sometimes, the arithmetic is shorter with one rather than the other.

ILLUSTRATION

$$27^{\frac{2}{3}} = 3^2 = 9 \text{ or } 27^{\frac{2}{3}} = 729^{\frac{1}{3}} = 9.$$

Note:
It may be shown that all of the standard laws of indices may be used for fractional indices.

1.3.2 RADICALS (or Surds)

The symbol "$\sqrt{\ }$" is called a **"radical"** (or **"surd"**). It is used to indicate the positive or **"principal"** square root of a number. Thus $\sqrt{16} = 4$ and $\sqrt{25} = 5$.

The number under the radical is called the **"radicand"**.

Most of our work on radicals will deal with square roots, but we may have occasion to use other roots of a number. For instance the **principal n-th root** of a number a is denoted by $\sqrt[n]{a}$, and is a number x such that $x^n = a$. The number n is called the **index** of the radical but, of course, when $n = 2$ we usually leave the index out.

ILLUSTRATIONS

1. $\sqrt[3]{64} = 4$ since $4^3 = 64$.
2. $\sqrt[3]{-64} = -4$ since $(-4)^3 = -64$.
3. $\sqrt[4]{81} = 3$ since $3^4 = 81$.
4. $\sqrt[5]{32} = 2$ since $2^5 = 32$.
5. $\sqrt[5]{-32} = -2$ since $(-2)^5 = -32$.

Note:
If the index of the radical is an odd number, then the radicand may be positive or negative; but if the index of the radical is an even number, then the radicand may not be negative since no even power of a negative number will ever give a negative result.

(a) Rules for Square Roots

In preparation for work which will follow in the next section, we list here the standard rules for square roots:

(i) $(\sqrt{a})^2 = a$

(ii) $\sqrt{a^2} = |a|$

(iii) $\sqrt{ab} = \sqrt{a}\sqrt{b}$

(iv) $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$

assuming that all of the radicals can be evaluated.

ILLUSTRATIONS

1. $\sqrt{9 \times 4} = \sqrt{36} = 6$ and $\sqrt{9} \times \sqrt{4} = 3 \times 2 = 6$.
2. $\sqrt{\frac{144}{36}} = \sqrt{4} = 2$ and $\frac{\sqrt{144}}{\sqrt{36}} = \frac{12}{6} = 2$.

(b) Rationalisation of Radical (or Surd) Expressions.

It is often desirable to eliminate expressions containing radicals from the denominator of a quotient. This process is called

rationalising the denominator.

The process involves multiplying numerator and denominator of the quotient by the same amount - an amount which eliminates the radicals in the denominator (often using the fact that the square root of a number multiplied by itself gives just the number;
i.e. $\sqrt{a}.\sqrt{a} = a$). We illustrate with examples:

EXAMPLES

1. Rationalise the surd form $\frac{5}{4\sqrt{3}}$

   Solution

   We simply multiply numerator and denominator by $\sqrt{3}$ to give

   $$\frac{5}{4\sqrt{3}} = \frac{5}{4\sqrt{3}} \times \frac{\sqrt{3}}{\sqrt{3}} = \frac{5\sqrt{3}}{12}.$$

2. Rationalise the surd form $\frac{\sqrt[3]{a}}{\sqrt[3]{b}}$

   Solution

   Here we observe that, if we can convert the denominator into the cube root of b^n, where n is a whole multiple of 3, then the square root sign will disappear.

   We have

   $$\frac{\sqrt[3]{a}}{\sqrt[3]{b}} = \frac{\sqrt[3]{a}}{\sqrt[3]{b}} \times \frac{\sqrt[3]{b^2}}{\sqrt[3]{b^2}} = \frac{\sqrt[3]{ab^2}}{\sqrt[3]{b^3}} = \frac{\sqrt[3]{ab^2}}{b}.$$

   If the denominator is of the form $\sqrt{a} + \sqrt{b}$, we multiply the numerator and the denominator by the expression $\sqrt{a} - \sqrt{b}$ because

   $$(\sqrt{a} + \sqrt{b})(\sqrt{a} - \sqrt{b}) = a - b.$$

3. Rationalise the surd form $\frac{4}{\sqrt{5}+\sqrt{2}}$.

   Solution

   Multiplying numerator and denominator by $\sqrt{5} - \sqrt{2}$ gives

   $$\frac{4}{\sqrt{5} + \sqrt{2}} \times \frac{\sqrt{5} - \sqrt{2}}{\sqrt{5} - \sqrt{2}} = \frac{4\sqrt{5} - 4\sqrt{2}}{3}.$$

4. Rationalise the surd form $\frac{1}{\sqrt{3}-1}$.

   Solution

   Multiplying numerator and denominator by $\sqrt{3}+1$ gives

$$\frac{1}{\sqrt{3}-1} \times \frac{\sqrt{3}+1}{\sqrt{3}+1} = \frac{\sqrt{3}+1}{2}.$$

(c) Changing numbers to and from radical form

The modulus of any number of the form $a^{\frac{m}{n}}$ can be regarded as the principal n-th root of a^m; i.e.

$$\mid a^{\frac{m}{n}} \mid = \sqrt[n]{a^m}.$$

If a number of the type shown on the left is converted to the type on the right, we are said to have expressed it in radical form.

If a number of the type on the right is converted to the type on the left, we are said to have expressed it in exponential form.

Note:
The word **"exponent"** is just another word for **"power"** or **"index"** and the standard rules of indices will need to be used in questions of the type discussed here.

EXAMPLES

1. Express the number $x^{\frac{2}{5}}$ in radical form.

   Solution

   The answer is just

$$\sqrt[5]{x^2}.$$

2. Express the number $\sqrt[3]{a^5b^4}$ in exponential form.

   Solution

   Here we have

$$\sqrt[3]{a^5b^4} = (a^5b^4)^{\frac{1}{3}} = a^{\frac{5}{3}}b^{\frac{4}{3}}.$$

1.3.3 EXERCISES

1. Simplify
   (a) $5^7 \times 5^{13}$; (b) $9^8 \times 9^5$; (c) $11^2 \times 11^3 \times 11^4$.
2. Simplify
   (a) $\frac{15^3}{15^2}$; (b) $\frac{4^{18}}{4^9}$; (c) $\frac{5^{20}}{5^{19}}$.
3. Simplify
   (a) a^7a^3; (b) a^4a^5;
   (c) $b^{11}b^{10}b$; (d) $3x^6 \times 5x^9$.
4. Simplify
   (a) $(7^3)^2$; (b) $(4^2)^8$; (c) $(7^9)^2$.
5. Simplify
   (a) $(x^2y^3)(x^3y^2)$; (b) $(2x^2)(3x^4)$;
   (c) $(a^2bc^2)(b^2ca)$; (d) $\frac{6c^2d^3}{3cd^2}$.
6. Simplify
   (a) $\left(4^{-3}\right)^2$ (b) $a^{13}a^{-2}$;
   (c) $x^{-9}x^{-7}$; (d) $x^{-21}x^2x$;
   (e) $\frac{x^2y^{-1}}{z^3} \div \frac{z^2}{x^{-1}y^3}$.
7. Without using a calculator, evaluate the following:
   (a) $\frac{4^{-8}}{4^{-6}}$; (b) $\frac{3^{-5}}{3^{-8}}$.
8. Evaluate the following:
   (a) $64^{\frac{1}{3}}$; (b) $144^{\frac{1}{2}}$;
   (c) $16^{-\frac{1}{4}}$; (d) $25^{-\frac{1}{2}}$;
   (e) $16^{\frac{3}{2}}$; (f) $125^{-\frac{2}{3}}$.
9. Simplify the following radicals:
   (a) $-\sqrt[3]{-8}$; (b) $\sqrt{36x^4}$; (c) $\sqrt{\frac{9a^2}{36b^2}}$.
10. Rationalise the following surd forms:
    (a) $\frac{\sqrt{2}}{\sqrt{3}}$; (b) $\frac{\sqrt[3]{18}}{\sqrt[3]{2}}$; (c) $\frac{2+\sqrt{5}}{\sqrt{3}-2}$; (d) $\frac{\sqrt{a}}{\sqrt{a}+3\sqrt{b}}$.
11. Change the following to exponential form:
    (a) $\sqrt[4]{7^2}$; (b) $\sqrt[5]{a^2b}$; (c) $\sqrt[3]{9^5}$.

12. Change the following to radical form:
 (a) $b^{\frac{3}{5}}$; (b) $r^{\frac{5}{3}}$; (c) $s^{\frac{7}{3}}$.

1.3.4 ANSWERS TO EXERCISES

1. (a) 5^{20}; (b) 9^{13}; (c) 11^{9}.
2. (a) 15; (b) 4^{9}; (c) 5.
3. (a) a^{10}; (b) a^{9}; (c) b^{22}; (d) $15x^{15}$.
4. (a) 7^{6}; (b) 4^{16}; (c) 7^{18}.
5. (a) x^5y^5; (b) $6x^6$; (c) $a^3b^3c^3$; (d) $2cd$.
6. (a) 4^{-6}; (b) a^{11}; (c) x^{-16}; (d) x^{-18}; (e) xy^2z^{-5}.
7. (a) $\frac{1}{16}$; (b) 27.
8. (a) 4; (b) ± 12; (c) $\pm\frac{1}{2}$;
 (d) $\pm\frac{1}{5}$; (e) ± 64; (f) $\frac{1}{25}$;
9. (a) 2; (b) $6x^2$; (c) $\left|\frac{a}{2b}\right|$.
10. (a) $\frac{\sqrt{6}}{3}$; (b) $\frac{\sqrt[3]{72}}{2} = \sqrt[3]{9}$; (c) $-(2+\sqrt{5})(2+\sqrt{3})$; (d) $\frac{a-3\sqrt{ab}}{a-9b}$.
11. (a) $\left|7^{\frac{1}{2}}\right|$; (b) $a^{\frac{2}{5}}b^{\frac{1}{5}}$; (c) $9^{\frac{5}{3}}$.
12. (a) $\sqrt[5]{b^3}$; (b) $\sqrt[3]{r^5}$; (c) $\sqrt[3]{s^7}$.

"JUST THE MATHS"

UNIT NUMBER

1.4

ALGEBRA 4
(Logarithms)

by

A.J.Hobson

UNIT 1.4 - ALGEBRA 4 - LOGARITHMS

1.4.1 COMMON LOGARITHMS

The system of numbers with which we normally count and calculate has a base of 10; this means that each of the successive digits of a particular number correspond to that digit multiplied by a certain power of 10.

For example

$$73,520 = 7 \times 10^4 + 3 \times 10^3 + 5 \times 10^2 + 2 \times 10^1.$$

Note:
Other systems (not discussed here) are sometimes used - such as the binary system which uses successive powers of 2.

The question now arises as to whether a given number can be expressed as a single power of 10, not necessarily an integer power. It will certainly need to be a **positive** number since powers of 10 are not normally negative (or even zero).

It can easily be verified by calculator, for instance that

$$1.99526 \simeq 10^{0.3}$$

and

$$2 \simeq 10^{0.30103}.$$

DEFINITION
In general, when it occurs that

$$x = 10^y,$$

for some positive number x, we say that y is the **"logarithm to base 10"** of x (or " **common logarithm**" of x) and we write

$$y = \log_{10} x.$$

EXAMPLES

1. $\log_{10} 1.99526 = 0.3$ from the illustrations above.
2. $\log_{10} 2 = 0.30103$ from the illustrations above.
3. $\log_{10} 1 = 0$ simply because $10^0 = 1$.

1.4.2 LOGARITHMS IN GENERAL

In practice, with scientific work, only two bases of logarithms are ever used; but it will be useful to include here a general discussion of the definition and properties of logarithms to **any** base so that unnecessary repetition may be avoided. We consider only positive bases of logarithms in the general discussion.

DEFINITION
If B is a fixed positive number and x is another positive number such that

$$x = B^y,$$

we say that y is the **"logarithm to base B"** of x and we write

$$y = \log_B x.$$

EXAMPLES

1. $\log_B 1 = 0$ simply because $B^0 = 1$.
2. $\log_B B = 1$ simply because $B^1 = B$.
3. $\log_B 0$ doesn't really exist because no power of B could ever be equal to zero. But, since a very large negative power of B will be a very small positive number, we usually write

$$\log_B 0 = -\infty.$$

1.4.3 USEFUL RESULTS

In preparation for the general properties of logarithms, we note the following two results which can be obtained directly from the definition of a logarithm:

(a) For any positive number x,

$$x = B^{\log_B x}.$$

In other words, any positive number can be expressed as a power of B without necessarily using a calculator.

We have simply replaced the y in the statement $x = B^y$ by $\log_B x$ in the equivalent statement $y = \log_B x$.

(b) For any number y,

$$y = \log_B B^y.$$

In other words, any number can be expressed in the form of a logarithm without necessarily using a calculator.

We have simply replaced x in the statement $y = \log_B x$ by B^y in the equivalent statement $x = B^y$.

1.4.4 PROPERTIES OF LOGARITHMS

The following properties were once necessary for performing numerical calculations before electronic calculators came into use. We do not use logarithms for this purpose nowadays; but we do need their properties for various topics in scientific mathematics.

(a) The Logarithm of Product.

$$\log_B p.q = \log_B p + \log_B q.$$

Proof:
We need to show that, when $p.q$ is expressed as a power of B, that power is the expression on the right hand side of the above formula.

From Result (a) of the previous section,

$$p.q = B^{\log_B p}.B^{\log_B q} = B^{\log_B p + \log_B q},$$

by elementary properties of indices.

The result therefore follows.

(b) The Logarithm of a Quotient

$$\log_B \frac{p}{q} = \log_B p - \log_B q.$$

Proof:
The proof is along similar lines to that in (i).

From Result (a) of the previous section,

$$\frac{p}{q} = \frac{B^{\log_B p}}{B^{\log_B q}} = B^{\log_B p - \log_B q},$$

by elementary properties of indices.

The result therefore follows.

(c) The Logarithm of an Exponential

$$\log_B p^n = n \log_B p,$$

where n need not be an integer.

Proof:
From Result (a) of the previous section,

$$p^n = \left(B^{\log_B p}\right)^n = B^{n \log_B p},$$

by elementary properties of indices.

(d) The Logarithm of a Reciprocal

$$\log_B \frac{1}{q} = -\log_B q.$$

Proof:
This property may be proved in two ways as follows:

Method 1.
The left-hand side $= \log_B 1 - \log_B q = 0 - \log_B q = -\log_B q$.

Method 2.
The left-hand side $= \log_B q^{-1} = -\log_B q$.

(e) Change of Base

$$\log_B x = \frac{\log_A x}{\log_A B}.$$

Proof:
Suppose $y = \log_B x$, then $x = B^y$ and hence

$$\log_A x = \log_A B^y = y \log_A B.$$

Thus,

$$y = \frac{\log_A x}{\log_A B}$$

and the result follows.

Note:
The result shows that the logarithms of any set of numbers to a given base will be directly

proportional to the logarithms of the same set of numbers to another given base. This is simply because the number $\log_A B$ is a constant.

1.4.5 NATURAL LOGARITHMS

It was mentioned earlier that, in scientific work, only two bases of logarithms are ever used. One of these is base 10 and the other is a base which arises **naturally** out of elementary calculus when discussing the simplest available result for the "derivative" (rate of change) of a logarithm.

This other base turns out to be a non-recurring, non-terminating decimal quantity (irrational number) which is equal to 2.71828.....and clearly this would be inconvenient to write into the logarithm notation.

We therefore denote it by e to give the **"natural logarithm"** of a number, x, in the form $\log_e x$, although most scientific books use the alternative notation $\ln x$.

Note:
From the earlier change of base formula we can say that

$$\log_{10} x = \frac{\log_e x}{\log_e 10} \quad \text{and} \quad \log_e x = \frac{\log_{10} x}{\log_{10} e}.$$

EXAMPLES

1. Solve for x the indicial equation

$$4^{3x-2} = 26^{x+1}.$$

Solution
The secret of solving an equation where an unknown quantity appears in a power (or index or exponent) is to take logarithms of both sides first.
Here we obtain

$$(3x - 2)\log_{10} 4 = (x + 1)\log_{10} 26;$$

$$(3x - 2)0.6021 = (x + 1)1.4150;$$

$$1.8063x - 1.2042 = 1.4150x + 1.4150;$$

$$(1.8603 - 1.4150)x = 1.4150 + 1.2042;$$

$$0.3913x = 2.6192;$$

$$x = \frac{2.6192}{0.3913} \simeq 6.6936$$

2. Rewrite the expression

$$4x + \log_{10}(x+1) - \log_{10} x - \frac{1}{2}\log_{10}(x^3 + 2x^2 - x)$$

as the common logarithm of a single mathematical expression.

Solution

The secret here is to make sure that every term in the given expression is converted, where necessary, to a logarithm with no multiple in front of it or behind it. In this case, we need first to write $4x = \log_{10} 10^{4x}$ and $\frac{1}{2}\log_{10}(x^3 + 2x^2 - x) = \log_{10}\left(x^3 + 2x^2 - x\right)^{\frac{1}{2}}$.

We can then use the results for the logarithms of a product and a quotient to give

$$\log_{10} \frac{10^{4x}(x+1)}{x\sqrt{(x^3 + 2x^2 - x)}}.$$

3. Rewrite without logarithms the equation

$$2x + \ln x = \ln(x-7).$$

Solution

This time, we need to convert both sides to the natural logarithm of a single mathematical expression in order to remove the logarithms completely.

$$2x + \ln x = \ln e^{2x} + \ln x = \ln xe^{2x}.$$

Hence,

$$xe^{2x} = x - 7.$$

4. Solve for x the equation

$$6\ln 4 + \ln 2 = 3 + \ln x.$$

Solution

In view of the facts that $6\ln 4 = \ln 4^6$ and $3 = \ln e^3$, the equation can be written

$$\ln 2(4^6) = \ln xe^3.$$

Hence,

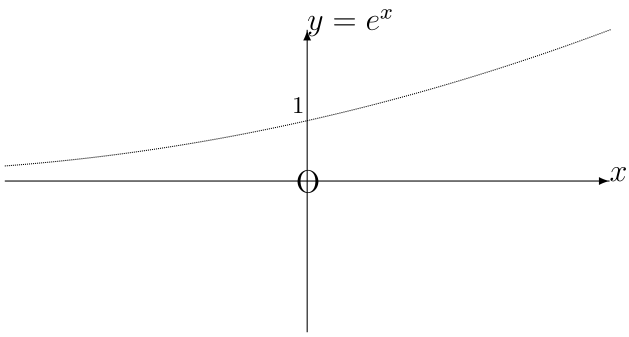
$$2(4^6) = xe^3,$$

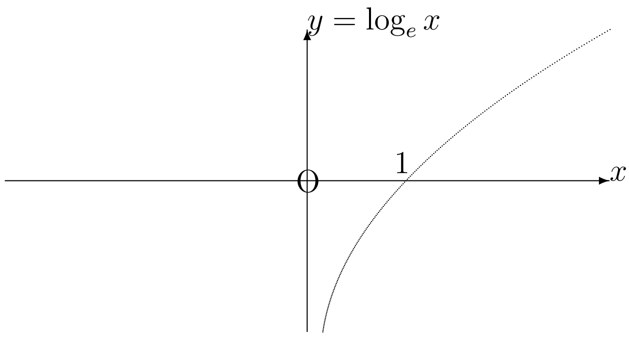
so that

$$x = \frac{2(4^6)}{e^3} \simeq 407.856$$

1.4.6 GRAPHS OF LOGARITHMIC AND EXPONENTIAL FUNCTIONS

In the applications of mathematics to science and engineering, two commonly used "functions" are $y = e^x$ and $y = \log_e x$. Their graphs are as follows:





They are closely linked with each other by virtue of the two equivalent statements:

$$P = \log_e Q \quad \text{and} \quad Q = e^P$$

for any number P and any <u>postive</u> number Q.

Because of these statements, we would expect similarities in the graphs of the equations

$$y = \log_e x \quad \text{and} \quad y = e^x.$$

1.4.7 LOGARITHMIC SCALES

In a certain kind of graphical work (see Unit 5.3), some use is made of a linear scale along which numbers can be allocated according to their logarithmic distances from a chosen origin of measurement.

Considering firstly that 10 is the base of logarithms, the number 1 is always placed at the zero of measurement (since $\log_{10} 1 = 0$); 10 is placed at the first unit of measurement (since $\log_{10} 10 = 1$), 100 is placed at the second unit of measurement (since $\log_{10} 100 = 2$), and so on.

Negative powers of 10 such as $10^{-1} = 0.1$, $10^{-2} = 0.01$ etc. are placed at the points corresponding to -1 and -2 etc. respectively on an ordinary linear scale.

The logarithmic scale appears therefore in **"cycles"**, each cycle corresponding to a range of numbers between two consecutive powers of 10.

Intermediate numbers are placed at intervals which correspond to their logarithm values.

An extract from a typical logarithmic scale would be as follows:

0.1 0.2 0.3 0.4 1 2 3 4 10

Notes:
(i) A given set of numbers will determine how many cycles are required on the logarithmic scale. For example .3, .6, 5, 9, 23, 42, 166 will require **four** cycles.

(ii) Commercially printed logarithmic scales do not specify the base of logarithms; the change of base formula implies that logarithms to different bases are proportional to each other and hence their logarithmic scales will have the same relative shape.

1.4.8 EXERCISES

1. Without using tables or a calculator, evaluate

   (a) $\log_{10} 27 \div \log_{10} 3$;
   (b) $(\log_{10} 16 - \log_{10} 2) \div \log_{10} 2$.

2. Using properties of logarithms where possible, solve for x the following equations:

   (a) $\log_{10} \frac{7}{2} + 2\log_{10} \frac{4}{3} - \log_{10} \frac{7}{45} = 1 + \log_{10} x$;
   (b) $2\log_{10} 6 - (\log_{10} 4 + \log_{10} 9) = \log_{10} x$.
   (c) $10^x = 5(2^{10})$.

3. From the definition of a logarithm or the change of base formula, evaluate the following:

   (a) $\log_2 7$;
   (b) $\log_3 0.04$;
   (c) $\log_5 3$;
   (d) $3\log_3 2 - \log_3 4 + \log_3 \frac{1}{2}$.

4. Obtain y in terms of x for the following equations:

   (a) $2\ln y = 1 - x^2$;
   (b) $\ln x = 5 - 3\ln y$.

5. Rewrite the following statements without logarithms:

   (a) $\ln x = -\frac{1}{2}\ln(1 - 2v^3) + \ln C$;
   (b) $\ln(1 + y) = \frac{1}{2}x^2 + \ln 4$;
   (c) $\ln(4 + y^2) = 2\ln(x + 1) + \ln A$.

6. (a) If $\frac{I_0}{I} = 10^{ac}$, find c in terms of the other quantities in the formula.
   (b) If $y^p = Cx^q$, find q in terms of the other quantities in the formula.

1.4.9 ANSWERS TO EXERCISES

1. (a) 3; (b) 3.
2. (a) 4; (b) 1; (c) 3.70927
3. (a) 2.807; (b) −2.930; (c) 0.683; (d) 0
4. (a)

$$y = e^{\frac{1}{2}(1-x^2)};$$

(b)

$$y = \sqrt[3]{\frac{e^5}{x}}.$$

5. (a)

$$x = \frac{C}{\sqrt{1-2v^3}};$$

(b)

$$1 + y = 4e^{\frac{x^2}{2}};$$

(c)

$$4 + y^2 = A(x+1)^2.$$

6. (a)

$$c = \frac{1}{a}\log_{10}\frac{I_0}{I};$$

(b)

$$q = \frac{p\log y - \log C}{\log x},$$

using any base.

"JUST THE MATHS"

UNIT NUMBER

1.5

ALGEBRA 5
(Manipulation of algebraic expressions)

by

A.J.Hobson

UNIT 1.5 - ALGEBRA 5
MANIPULATION OF ALGEBRAIC EXPRESSIONS

1.5.1 SIMPLIFICATION OF EXPRESSIONS

An algebraic expression will, in general, contain a mixture of alphabetical symbols together with one or more numerical quantities; some of these symbols and numbers may be bracketted together.

Using the Language of Algebra and the Laws of Algebra discussed earlier, the method of simplification is to remove brackets and collect together any terms which have the same format

Some elementary illustrations are as follows:

1. $a + a + a + 3 + b + b + b + b + 8 \equiv 3a + 4b + 11$.
2. $11p^2 + 5q^7 - 8p^2 + q^7 \equiv 3p^2 + 6q^7$.
3. $a(2a - b) + b(a + 5b) - a^2 - 4b^2 \equiv 2a^2 - ab + ba + 5b^2 - a^2 - 4b^2 \equiv a^2 + b^2$.

More frequently, the expressions to be simplified will involve symbols which represent both the constants and variables encountered in scientific work. Typical examples in pure mathematics use symbols like x, y and z to represent the variable quantities.

Further illustrations use this kind of notation and, for simplicity, we shall omit the full-stop type of multiplication sign between symbols.

1. $x(2x + 5) + x^2(3 - x) \equiv 2x^2 + 5x + 3x^2 - x^3 \equiv 5x^2 + 5x - x^3$.
2. $x^{-1}(4x - x^2) - 6(1 - 3x) \equiv 4 - x - 6 + 18x \equiv 17x - 2$.

We need also to consider the kind of expression which involves two or more brackets multiplied together; but the routine is just an extension of what has already been discussed.

For example consider the expression

$$(a + b)(c + d).$$

Taking the first bracket as a single item for the moment, the Distributive Law gives

$$(a + b)c + (a + b)d.$$

Using the Distributive Law a second time gives

$$ac + bc + ad + bd.$$

In other words, each of the two terms in the first bracket are multiplied by each of the two terms in the second bracket, giving four terms in all.

Again, we illustrate with examples:

EXAMPLES

1. $(x + 3)(x - 5) \equiv x^2 + 3x - 5x - 15 \equiv x^2 - 2x - 15.$
2. $(x^3 - x)(x + 5) \equiv x^4 - x^2 + 5x^3 - 5x.$
3. $(x + a)^2 \equiv (x + a)(x + a) \equiv x^2 + ax + ax + a^2 \equiv x^2 + 2ax + a^2.$
4. $(x + a)(x - a) \equiv x^2 + ax - ax - a^2 \equiv x^2 - a^2.$

The last two illustrations above are significant for later work because they incorporate, respectively, the standard results for a **"perfect square"** and **"the difference between two squares"**.

1.5.2 FACTORISATION

Introduction
In an algebraic context, the word **"factor"** means the same as **"multiplier"**. Thus, to factorise an algebraic expression, is to write it as a product of separate multipliers or factors.

Some simple examples will serve to introduce the idea:

EXAMPLES

1.
$$3x + 12 \equiv 3(x + 4).$$

2.
$$8x^2 - 12x \equiv x(8x - 12) \equiv 4x(2x - 3).$$

3.
$$5x^2 + 15x^3 \equiv x^2(5 + 15x) \equiv 5x^2(1 + 3x).$$

4. $$6x + 3x^2 + 9xy \equiv x(6 + 3x + 9y) \equiv 3x(2 + x + 3y).$$

Note:
When none of the factors can be broken down into simpler factors, the original expression is said to have been factorised into **"irreducible factors"**.

Factorisation of quadratic expressions

A **"quadratic expression"** is an expression of the form

$$ax^2 + bx + c,$$

where, usually, a, b and c are fixed numbers (constants) while x is a variable number. The numbers a and b are called, respectively, the **"coefficients"** of x^2 and x while c is called the **"constant term"**; but, for brevity, we often say that the quadratic expression has coefficients a, b and c.

Note:
It is important that the coefficient a does not have the value zero otherwise the expression is not quadratic but **"linear"**.

The method of factorisation is illustrated by examples:

(a) When the coefficient of x^2 is 1

EXAMPLES

1. $$x^2 + 5x + 6 \equiv (x + m)(x + n) \equiv x^2 + (m + n)x + mn.$$

   This implies that $5 = m + n$ and $6 = mn$ which, by inspection gives $m = 2$ and $n = 3$. Hence
   $$x^2 + 5x + 6 \equiv (x + 2)(x + 3).$$

2. $$x^2 + 4x - 21 \equiv (x + m)(x + n) \equiv x^2 + (m + n)x + mn.$$

   This implies that $4 = m + n$ and $-21 = mn$ which, by inspection, gives $m = -3$ and $n = 7$. Hence
   $$x^2 + 4x - 21 \equiv (x - 3)(x + 7).$$

Notes:
(i) In general, for simple cases, it is better to try to carry out the factorisation entirely by inspection. This avoids the cumbersome use of m and n in the above two examples as follows:

$$x^2 + 2x - 8 \equiv (x+?)(x+?).$$

The two missing numbers must be such that their sum is 2 and their product is -8. The required values are therefore -2 and 4. Hence

$$x^2 + 2x - 8 \equiv (x - 2)(x + 4).$$

(ii) It is necessary, when factorising a quadratic expression, to be aware that either a perfect square or the difference of two squares might be involved. In these cases, the factorisation is a little simpler. For instance:

$$x^2 + 10x + 25 \equiv (x + 5)^2$$

and

$$x^2 - 64 \equiv (x - 8)(x + 8).$$

(iii) Some quadratic expressions will not conveniently factorise at all. For example, in the expression

$$x^2 - 13x + 2,$$

we cannot find two whole numbers whose sum is -13 while, at the same time, their product is 2.

(b) When the coefficient of x^2 is not 1

Quadratic expressions of this kind are usually more difficult to factorise than those in the previous paragraph. We first need to determine the possible pairs of factors of the coefficient of x^2 and the possible pairs of factors of the constant term; then we need to consider the possible combinations of these which provide the correct factors of the quadratic expression.

EXAMPLES

1. To factorise the expression

$$2x^2 + 11x + 12,$$

we observe that 2 is the product of 2 and 1 only, while 12 is the product of either 12 and 1, 6 and 2 or 4 and 3. All terms of the quadratic expression are positive and hence we may try $(2x+1)(x+12)$, $(2x+12)(x+1)$, $(2x+6)(x+2)$, $(2x+2)(x+6)$, $(2x+4)(x+3)$ and $(2x+3)(x+4)$. Only the last of these is correct and so

$$2x^2 + 11x + 12 \equiv (2x+3)(x+4).$$

2. To factorise the expression

$$6x^2 + 7x - 3,$$

we observe that 6 is the product of either 6 and 1 or 3 and 2 while 3 is the product of 3 and 1 only. A negative constant term implies that, in the final result, its two factors must have opposite signs. Hence we may try $(6x+3)(x-1)$, $(6x-3)(x+1)$, $(6x+1)(x-3)$, $(6x-1)(x+3)$, $(3x+3)(2x-1)$, $(3x-3)(2x+1)$, $(3x+1)(2x-3)$ and $(3x-1)(2x+3)$. Again, only the last of these is correct and so

$$6x^2 + 7x - 3 \equiv (3x-1)(2x+3).$$

Note:
The more factors there are in the coefficients considered, the more possibilities there are to try of the final factorisation.

1.5.3 COMPLETING THE SQUARE IN A QUADRATIC EXPRESSION

The following work is based on the standard expansions

$$(x+a)^2 \equiv x^2 + 2ax + a^2$$

and

$$(x-a)^2 \equiv x^2 - 2ax + a^2.$$

Both of these last expressions are called **"complete squares"** (or **"perfect squares"**) in which we observe that one half of the coefficient of x, when multiplied by itself, gives the constant term. That is

$$(\frac{1}{2} \times 2a)^2 = a^2.$$

ILLUSTRATIONS

1.

$$x^2 + 6x + 9 \equiv (x+3)^2.$$

2.

$$x^2 - 8x + 16 \equiv (x-4)^2.$$

3.

$$4x^2 - 4x + 1 \equiv 4\left[x^2 - x + \frac{1}{4}\right] \equiv 4\left(x - \frac{1}{2}\right)^2.$$

Of course it may happen that a given quadratic expression is NOT a complete square; but, by using one half of the coefficient of x, we may express it as the sum or difference of a complete square and a constant. This process is called **"completing the square"**, and the following examples illustrate it:

EXAMPLES

1.

$$x^2 + 6x + 11 \equiv (x+3)^2 + 2.$$

2.

$$x^2 - 8x + 7 \equiv (x-4)^2 - 9.$$

3.

$$\begin{aligned}4x^2 - 4x + 5 &\equiv 4\left[x^2 - x + \frac{5}{4}\right]\\ &\equiv 4\left[\left(x - \frac{1}{2}\right)^2 - \frac{1}{4} + \frac{5}{4}\right]\\ &\equiv 4\left[\left(x - \frac{1}{2}\right)^2 + 1\right]\\ &\equiv 4\left(x - \frac{1}{2}\right)^2 + 4.\end{aligned}$$

1.5.4 ALGEBRAIC FRACTIONS

We first recall the basic rules for combining fractions, namely

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \quad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}, \quad \frac{a}{b} \div \frac{c}{d} = \frac{ad}{bc}.$$

We also note that a single algebraic fraction may sometimes be simplified by the cancellation of common factors between the numerator and the denominator.

EXAMPLES

1.
$$\frac{5}{25+15x} \equiv \frac{1}{5+3x}, \quad \text{assuming that} \quad x \neq -\frac{5}{3}.$$

2.
$$\frac{4x}{3x^2+x} \equiv \frac{4}{3x+1}, \quad \text{assuming that} \quad x \neq 0 \quad \text{or} \quad -\frac{1}{3}.$$

3.
$$\frac{x+2}{x^2+3x+2} \equiv \frac{x+2}{(x+2)(x+1)} \equiv \frac{1}{x+1}, \quad \text{assuming that} \quad x \neq -1 \quad \text{or} \quad -2.$$

These elementary principles may now be used with more advanced combinations of algebraic fractions

EXAMPLES

1. Simplify the expression
$$\frac{3x+6}{x^2+3x+2} \times \frac{x+1}{2x+8}.$$

   Solution
   Using factorisation where possible, together with the rule for multiplying fractions, we obtain
$$\frac{3(x+2)(x+1)}{2(x+4)(x+1)(x+2)} \equiv \frac{3}{2(x+4)},$$
   assuming that $x \neq -1, -2$ or -4.

2. Simplify the expression
$$\frac{3}{x+2} \div \frac{x}{2x+4}.$$

   Solution
   Using factorisation where possible together with the rule for dividing fractions, we obtain
$$\frac{3}{x+2} \times \frac{2x+4}{x} \equiv \frac{3}{x+2} \times \frac{2(x+2)}{x} \equiv \frac{6}{x},$$
   assuming that $x \neq 0$ or -2.

3. Express
$$\frac{4}{x+y} - \frac{3}{y}$$
as a single fraction.
Solution
From the basic rule for adding and subtracting fractions, we obtain
$$\frac{4y - 3(x+y)}{(x+y)y} \equiv \frac{y-3x}{(x+y)y},$$
assuming that $y \neq 0$ and $x \neq -y$.

4. Express
$$\frac{x}{x+1} + \frac{4-x^2}{x^2-x-2}$$
as a single fraction.
Solution
This example could be tackled in the same way as the previous one but it is worth noticing that $x^2 - x - 2 \equiv (x+1)(x-2)$. Consequently, it is worth putting both fractions over the simplest common denominator, namely $(x+1)(x-2)$. Hence we obtain, if $x \neq 2$ or -1,
$$\frac{x(x-2)}{(x+1)(x-2)} + \frac{4-x^2}{(x+1)(x-2)} \equiv \frac{x^2-2x+4-x^2}{(x+1)(x-2)} \equiv \frac{2(2-x)}{(x+1)(x-2)} \equiv -\frac{2}{x+1}.$$

1.5.5 EXERCISES

1. Write down in their simplest forms
(a) $5a - 2b - 3a + 6b$; (b) $11p + 5q - 2q + p$.
2. Simplify the following expressions:
(a) $3x^2 - 2x + 5 - x^2 + 7x - 2$; (b) $x^3 + 5x^2 - 2x + 1 + x - x^2$.
3. Expand and simplify the following expressions:
(a) $x(x^2 - 3x) + x^2(4x + 7)$; (b) $(2x-1)(2x+1) - x^2 + 5x$;
(c) $(x+3)(2x^2 - 5)$; (d) $2(3x+1)^2 + 5(x-7)^2$.
4. Factorise the following expressions:
(a) $xy + 4x^2y$; (b) $2abc - 6ab^2$;
(c) $\pi r^2 + 2\pi rh$; (d) $2xy^2z + 4x^2z$.

5. Factorise the following quadratic expressions:
   (a) $x^2 + 8x + 12$; (b) $x^2 + 11x + 18$; (c) $x^2 + 13x - 30$;
   (d) $3x^2 + 11x + 6$; (e) $4x^2 - 12x + 9$; (f) $9x^2 - 64$.

6. Complete the square in the following quadratic expressions:
   (a) $x^2 - 10x - 26$; (b) $x^2 - 5x + 4$; (c) $7x^2 - 2x + 1$.

7. Simplify the following:
   (a) $\frac{x^2+4x+4}{x^2+5x+6}$; (b) $\frac{x^2-1}{x^2+2x+1}$,
   assuming the values of x to be such that no denominators are zero.

8. Express each of the following as a single fraction:
   (a) $\frac{3}{x} + \frac{4}{y}$; (b) $\frac{4}{x} - \frac{6}{2x}$;
   (c) $\frac{1}{x+1} + \frac{1}{x+2}$; (d) $\frac{5x}{x^2+5x+4} - \frac{3}{x+4}$,
   assuming that the values of x and y are such that no denominators are zero.

1.5.6 ANSWERS TO EXERCISES

1. (a) $2a + 4b$; (b) $12p + 3q$.

2. (a) $2x^2 + 5x + 3$; (b) $x^3 + 4x^2 - x + 1$.

3. (a) $5x^3 + 4x^2$; (b) $3x^2 + 5x - 1$.
   (c) $2x^3 + 6x^2 - 5x - 15$; (d) $23x^2 - 58x + 247$.

4. (a) $xy(1 + 4x)$; (b) $2ab(c - 3b)$;
   (c) $\pi r(r + 2h)$; (d) $2xz(y^2 + 2x)$.

5. (a) $(x + 2)(x + 6)$; (b) $(x + 2)(x + 9)$; (c) $(x - 2)(x + 15$;
   (d) $(3x + 2)(x + 3)$; (e) $(2x - 3)^2$; (f) $(3x - 8)(3x + 8)$.

6. (a) $(x - 5)^2 - 51$; (b) $\left(x - \frac{5}{2}\right)^2 - \frac{9}{4}$; (c) $7\left[\left(x - \frac{1}{7}\right)^2 + \frac{6}{49}\right]$.

7. (a) $\frac{x+2}{x+3}$; (b) $\frac{x-1}{x+1}$.

8. (a) $\frac{3y+4x}{xy}$; (b) $\frac{1}{x}$;
   (c) $\frac{2x+3}{(x+1)(x+2)}$; (d) $\frac{2x-3}{x^2+5x+4}$.

"JUST THE MATHS"

UNIT NUMBER

1.6

ALGEBRA 6
(Formulae and algebraic equations)

by

A.J.Hobson

UNIT 1.6 - ALGEBRA 6 - FORMULAE AND ALGEBRAIC EQUATIONS

1.6.1 TRANSPOSITION OF FORMULAE

In dealing with technical formulae, it is often required to single out one of the quantities involved in terms of all the others. We are said to **"transpose the formula"** and make that quantity **"the subject of the equation"**.

In order to do this, steps of the following types may be carried out on both sides of a given formula:

(a) Addition or subtraction of the same value;

(b) Multiplication or division by the same value;

(c) The raising of both sides to equal powers;

(d) Taking logarithms of both sides.

EXAMPLES

1. Make x the subject of the formula

$$y = 3(x + 7).$$

Solution
Dividing both sides by 3 gives $\frac{y}{3} = x + 7$; then subtracting 7 gives $x = \frac{y}{3} - 7$.

2. Make y the subject of the formula

$$a = b + c\sqrt{x^2 - y^2}.$$

Solution
(i) Subtracting b gives $a - b = c\sqrt{x^2 - y^2}$;

(ii) Dividing by c gives $\frac{a-b}{c} = \sqrt{x^2 - y^2}$;

(iii) Squaring both sides gives $\left(\frac{a-b}{c}\right)^2 = x^2 - y^2$;

(iv) Subtracting x^2 gives $\left(\frac{a-b}{c}\right)^2 - x^2 = -y^2$;

(v) Multiplying throughout by -1 gives $x^2 - \left(\frac{a-b}{c}\right)^2 = y^2$;

(vi) Taking square roots of both sides gives

$$y = \pm\sqrt{x^2 - \left(\frac{a-b}{c}\right)^2}.$$

3. Make x the subject of the formula

$$e^{2x-1} = y^3.$$

Solution
Taking natural logarithms of both sides of the formula

$$2x - 1 = 3\ln y.$$

Hence

$$x = \frac{3\ln y + 1}{2}.$$

Note:
A genuine scientific formula will usually involve quantities which can assume only positive values; in which case we can ignore the negative value of a square root.

1.6.2 SOLUTION OF LINEAR EQUATIONS

A Linear Equation in a variable quantity x has the general form

$$ax + b = c.$$

Its solution is obtained by first subtracting b from both sides then dividing both sides by a. That is

$$x = \frac{c-b}{a}.$$

EXAMPLES

1. Solve the equation

$$5x + 11 = 20.$$

Solution
The solution is clearly $x = \frac{20-11}{5} = \frac{9}{5} = 1.8$

2. Solve the equation
$$3 - 7x = 12.$$

Solution
This time, the solution is $x = \frac{12-3}{-7} = \frac{9}{-7} \simeq -1.29$

1.6.3 SOLUTION OF QUADRATIC EQUATIONS

The standard form of a quadratic equation is
$$ax^2 + bx + c = 0,$$

where a, b and c are constants and $a \neq 0$.

We shall discuss three methods of solving such an equation related very closely to the previous discussion on quadratic expressions. The first two methods can be illustrated by examples.

(a) By Factorisation
This method depends on the ability to determine the factors of the left hand side of the given quadratic equation. This will usually be by trial and error.

EXAMPLES

1. Solve the quadratic equation
$$6x^2 + x - 2 = 0.$$

Solution
In factorised form, the equation can be written
$$(3x + 2)(2x - 1) = 0.$$
Hence, $x = -\frac{2}{3}$ or $x = \frac{1}{2}$.

2. Solve the quadratic equation
$$15x^2 - 17x - 4 = 0.$$

Solution
In factorised form, the equation can be written
$$(5x + 1)(3x - 4) = 0.$$
Hence, $x = -\frac{1}{5}$ or $x = \frac{4}{3}$

(b) By Completing the square

By looking at some numerical examples of this method, we shall be led naturally to a third method involving a **universal** formula for solving any quadratic equation.

EXAMPLES

1. Solve the quadratic equation
$$x^2 - 4x - 1 = 0.$$

   Solution
   On completing the square, the equation can be written
$$(x-2)^2 - 5 = 0.$$
   Thus,
$$x - 2 = \pm\sqrt{5}.$$
   That is,
$$x = 2 \pm \sqrt{5}.$$
   Left as it is, this is an answer in **"surd form"** but it could, of course, be expressed in decimals as 4.236 and −0.236.

2. Solve the quadratic equation
$$4x^2 + 4x - 2 = 0.$$

   Solution
   The equation may be written
$$4\left[x^2 + x - \frac{1}{2}\right] = 0$$
   and, on completing the square,
$$4\left[\left(x + \frac{1}{2}\right)^2 - \frac{3}{4}\right] = 0.$$
   Hence,
$$\left(x + \frac{1}{2}\right)^2 = \frac{3}{4},$$
   giving
$$x + \frac{1}{2} = \pm\sqrt{\frac{3}{4}}.$$

That is,

$$x = -\frac{1}{2} \pm \sqrt{\frac{3}{4}}$$

or

$$x = \frac{-1 \pm \sqrt{3}}{2}.$$

(c) By the Quadratic Formula

Starting now with an arbitrary quadratic equation

$$ax^2 + bx + c = 0,$$

we shall use the method of completing the square in order to establish the **general** solution.

The sequence of steps is as follows:

$$a\left[x^2 + \frac{b}{a}x + \frac{c}{a}\right] = 0;$$

$$a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right] = 0;$$

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a^2} - \frac{c}{a};$$

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2}{4a^2} - \frac{c}{a}};$$

$$x = -\frac{b}{2a} \pm \sqrt{\frac{b^2 - 4ac}{4a^2}};$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Note:
The quantity $b^2 - 4ac$ is called the **"discriminant"** of the equation and gives either two solutions, one solution or no solutions according as its value is positive, zero or negative.

The single solution case is usually interpreted as a pair of coincident solutions while the no solution case really means no **real** solutions. A more complete discussion of this case arises in the subject of **"complex numbers"** (see Unit 6.1).

EXAMPLES

Use the quadratic formula to solve the following:

1.

$$x^2 + 2x - 35 = 0.$$

Solution

$$x = \frac{-2 \pm \sqrt{4+140}}{2} = \frac{-2 \pm 12}{2} = 5 \quad \text{or} \quad -7.$$

2.

$$2x^2 - 3x - 7 = 0.$$

Solution

$$x = \frac{3 \pm \sqrt{9+56}}{4} = \frac{3 \pm \sqrt{65}}{4} = \frac{3 \pm 8.062}{4} \simeq 2.766 \quad \text{or} \quad -1.266$$

3.

$$9x^2 - 6x + 1 = 0.$$

Solution

$$x = \frac{6 \pm \sqrt{36-36}}{18} = \frac{6}{18} = \frac{1}{3} \quad \text{only}.$$

4.

$$5x^2 + x + 1 = 0.$$

Solution

$$x = \frac{-1 \pm \sqrt{1-20}}{10}.$$

Hence, there are no real solutions

1.6.4 EXERCISES

1. Make the given symbol the subject of the following formulae
   (a) x: $a(x-a) = b(x+b)$;
   (b) b: $a = \frac{2-7b}{3+5b}$;
   (c) r: $n = \frac{1}{2L}\sqrt{\frac{r}{p}}$;
   (d) x: $ye^{x^2+1} = 5$.
2. Solve, for x, the following equations
   (a) $14x = 35$;
   (b) $3x - 4.7 = 2.8$;
   (c) $4(2x-5) = 3(2x+8)$.
3. Solve the following quadratic equations by factorisation:
   (a) $x^2 + 5x - 14 = 0$;
   (b) $8x^2 + 2x - 3 = 0$.
4. Where possible, solve the following quadratic equations by the formula:
   (a) $2x^2 - 3x + 1 = 0$; (b) $4x = 45 - x^2$;
   (c) $16x^2 - 24x + 9 = 0$; (d) $3x^2 + 2x + 11 = 0$.

1.6.5 ANSWERS TO EXERCISES

1. (a) $x = \frac{b^2+a^2}{a-b}$;
   (b) $b = \frac{2-3a}{7+5a}$;
   (c) $r = 4n^2L^2p$;
   (d) $x = \pm\sqrt{\ln 5 - \ln y - 1}$.
2. (a) 2.5; (b) 2.5; (c) 22.
3. (a) $x = -7$, $x = 2$;
   (b) $x = -\frac{3}{4}$, $x = \frac{1}{2}$.
4. (a) $x = 1$, $x = \frac{1}{2}$;
   (b) $x = 5$, $x = -9$;
   (c) $x = \frac{3}{4}$ only;
   (d) No solutions.

"JUST THE MATHS"

UNIT NUMBER

1.7

ALGEBRA 7
(Simultaneous linear equations)

by

A.J.Hobson

UNIT 1.7 - ALGEBRA 7 - SIMULTANEOUS LINEAR EQUATIONS

Introduction

When Mathematics is applied to scientific work, it is often necessary to consider several statements involving several variables which are required to have a common solution for those variables. We illustrate here the case of two simultaneous linear equations in two variables x and y and three simultaneous linear equations in three unknowns x, y and z.

1.7.1 TWO SIMULTANEOUS LINEAR EQUATIONS IN TWO UNKNOWNS

Consider the simultaneous linear equations:

$$\begin{aligned} ax + by &= p, \\ cx + dy &= q. \end{aligned}$$

To obtain the solution we first eliminate one of the variables in order to calculate the other. For instance, to eliminate x, we try to make coefficient of x the same in both equations so that, subtracting one statement from the other, x disappears. In this case, we multiply the first equation by c and the second equation by a to give

$$\begin{aligned} cax + cby &= cp, \\ acx + ady &= aq. \end{aligned}$$

Subtracting the second equation from the first, we obtain

$$y(cb - ad) = cp - aq$$

which, in turn, means that

$$y = \frac{cp - aq}{cb - ad} \quad \text{provided} \quad cb - ad \neq 0.$$

Having found the value of y, we could then either substitute back into one of the original equations to find x or begin again by eliminating y.

However, it is better not to think of the above explanation as providing a **formula** for solving two simultaneous linear equations. Rather, a numerical example should be dealt with from first principles with the numbers provided.

Note:
$cb - ad = 0$ relates to a degenerate case in which the left hand sides of the two equations are proportional to each other. Such cases will not be dealt with at this stage.

EXAMPLE

Solve the simultaneous linear equations

$$\begin{aligned} 6x - 2y &= 1, \qquad (1) \\ 4x + 7y &= 9. \qquad (2) \end{aligned}$$

Solution

Multiplying the first equation by 4 and the second equation by 6,

$$\begin{aligned} 24x - 8y &= 4, \qquad (4) \\ 24x + 42y &= 54. \qquad (5) \end{aligned}$$

Subtracting the second of these from the first, we obtain $-50y = -50$ and hence $y = 1$.

Substituting back into equation (1), $6x - 2 = 1$, giving $6x = 3$, and, hence, $x = \frac{1}{2}$.

Alternative Method

Multiplying the first equation by 7 and the second equation by -2, we obtain

$$\begin{aligned} 42x - 14y &= 7, \qquad (5) \\ -8x - 14y &= -18. \qquad (6) \end{aligned}$$

Subtracting equation (6) from equation (5) gives $50x = 25$ and hence, $x = \frac{1}{2}$.

Substituting into equation (1) gives $3 - 2y = 1$ and hence, $-2y = -2$; that is $y = 1$.

1.7.2 THREE SIMULTANEOUS LINEAR EQUATIONS IN THREE UNKNOWNS

Here, we consider three simultaneous equations of the general form

$$\begin{aligned} a_1x + b_1y + c_1z &= k_1, \\ a_2x + b_2y + c_2z &= k_2, \\ a_3x + b_3y + c_3z &= k_3; \end{aligned}$$

but the method will be illustrated by a particular example.

The object of the method is to eliminate one of the variables from two different pairs of the three equations so that we are left with a pair of simultaneous equations from which to calculate the other two variables.

EXAMPLE

Solve, for x, y and z, the simultaneous linear equations

$$\begin{aligned} x - y + 2z &= 9, && (1) \\ 2x + y - z &= 1, && (2) \\ 3x - 2y + z &= 8. && (3) \end{aligned}$$

Solution

Firstly, we may eliminate z from equations (2) and (3) by adding them together. We obtain

$$5x - y = 9. \qquad (4)$$

Secondly, we may eliminate z from equations (1) and (2) by doubling equation (2) and adding it to equation (1). We obtain

$$5x + y = 11. \qquad (5)$$

If we now add equation (4) to equation (5), y will be eliminated to give

$$10x = 20 \quad \text{or} \quad x = 2.$$

Similarly, if we subtract equation (4) from equation (5), x will be eliminated to give

$$2y = 2 \quad \text{or} \quad y = 1.$$

Finally, if we substitute our values of x and y into one of the original equations [say equation (3)] we obtain

$$z = 8 - 3x + 2y = 8 - 6 + 2 = 4.$$

Thus,

$$x = 2, \quad y = 1 \quad \text{and} \quad z = 4.$$

1.7.3 ILL-CONDITIONED EQUATIONS

In the simultaneous linear equations of genuine scientific problems, the coefficients will often be decimal quantitities that have already been subjected to rounding errors; and the solving process will tend to amplify these errors. The result may be that such errors swamp the values of the variables being solved for; and we have what is called an **"ill-conditioned"** set of equations. The opposite of this is a **"well-conditioned"** set of equations and all of those so far discussed have been well-conditioned. But let us consider, now, the following example:

EXAMPLE

The simultaneous linear equations

$$\begin{aligned} x + y &= 1, \\ 1.001x + y &= 2 \end{aligned}$$

have the common solution $x = 1000$, $y = -999$.

However, suppose that the coefficient of x in the second equation is altered to 1.000, which is a mere 0.1%. Then the equations have no solution at all since $x + y$ cannot be equal to 1 and 2 at the same time.

Secondly, suppose that the coefficient of x in the second equation is altered to 0.999 which is still only a 0.2% alteration.

The solutions obtained are now $x = -1000$, $y = 1001$ and so a change of about 200% has occurred in original values of x and y.

1.7.4 EXERCISES

1. Solve, for x and y, the following pairs of simultaneous linear equations:

   (a)

$$\begin{aligned} x - 2y &= 5, \\ 3x + y &= 1; \end{aligned}$$

   (b)

$$\begin{aligned} 2x + 3y &= 42, \\ 5x - y &= 20. \end{aligned}$$

2. Solve, for x, y and z, the following sets of simultaneous equations:

   (a)

$$\begin{aligned} x + y + z &= 0, \\ 2x - y - 3z &= 4, \\ 3x + 3y &= 7; \end{aligned}$$

   (b)

$$\begin{aligned} x + y - 10 &= 0, \\ y + z - 3 &= 0, \\ x + z + 1 &= 0; \end{aligned}$$

   (c)

$$\begin{aligned} 2x - y - z &= 6, \\ x + 3y + 2z &= 1, \\ 3x - y - 5z &= 1; \end{aligned}$$

(d)

$$\begin{aligned} 2x - 5y + 2z &= 14, \\ 9x + 3y - 4z &= 13, \\ 7x + 3y - 2z &= 3; \end{aligned}$$

(e)

$$\begin{aligned} 4x - 7y + 6z &= -18, \\ 5x + y - 4z &= -9, \\ 3x - 2y + 3z &= 12. \end{aligned}$$

3. Solve the simultaneous linear equations

$$\begin{aligned} 1.985x - 1.358y &= 2.212, \\ 0.953x - 0.652y &= 1.062, \end{aligned}$$

and compare with the solutions obtained by changing the constant term, 1.062, of the second equation to 1.061.

1.7.5 ANSWERS TO EXERCISES

1. (a) $x = 1,\ y = -2$; (b) $x = 6,\ y = 10$.
2. (a) $x = -\frac{2}{9},\ y = \frac{23}{9},\ z = -\frac{7}{3}$;
   (b) $x = 3,\ y = 7,\ z = -4$;
   (c) $x = 3,\ y = -2,\ z = 2$;
   (d) $x = 1,\ y = -4,\ z = -4$;
   (e) $x = 3,\ y = 12,\ z = 9$.
3. $$x \simeq 0.6087, \quad y \simeq -0.7391$$

   compared with

   $$x \simeq 30.1304, \quad y \simeq 42.413$$

   a change of 4850% in x and 5839% in y.

"JUST THE MATHS"

UNIT NUMBER

1.8

ALGEBRA 8
(Polynomials)

by

A.J.Hobson

UNIT 1.8 - ALGEBRA 8 - POLYNOMIALS

Introduction

The work already covered in earlier units has frequently been concerned with mathematical expressions involving constant quantities together positive integer powers of a variable quantity (usually x). The general form of such expressions is

$$a_0 + a_1x + a_2x^2 + a_3x^3 + ... + a_nx^n$$

which is called a
"polynomial of degree n in x", having **"coefficients"** $a_0, a_1, a_2, a_3, ...a_n$, usually constant.

Note:
Polynomials of degree 1, 2 and 3 are called respectively **"linear"**, **"quadratic"** and **"cubic"** polynomials.

1.8.1 THE FACTOR THEOREM

If $P(x)$ denotes an algebraic polynomial which has the value zero when $x = \alpha$, then $x - \alpha$ is a factor of the polynomial and
$P(x) \equiv (x - \alpha)\times$ another polynomial, $Q(x)$, of one degree lower.

Notes:
(i) The statement of this Theorem includes some functional notation (i.e. $P(x), Q(x)$) which will be discussed fully as an introduction to the subject of Calculus in Unit 10.1.

(ii) $x = \alpha$ is called a **"root"** of the polynomial.

1.8.2 APPLICATION TO QUADRATIC AND CUBIC EXPRESSIONS

(a) Quadratic Expressions

Suppose we are given a quadratic expression where we suspect that at least one of the values of x making it zero is a whole number. We may systematically try

$$x = 0, 1, -1, 2, -2, 3, -3, 4, -4, ...$$

until such a value of x (say $x = \alpha$) is located. Then one of the factors of the quadratic expression is $x - \alpha$ enabling us to determine the other factor(s) easily.

EXAMPLES

1. By trial and error, the quadratic expression
$$x^2 + 2x - 3$$
has the value zero when $x = 1$; hence, $x - 1$ is a factor.
By further trial and error, the complete factorisation is
$$(x - 1)(x + 3).$$

2. By trial and error, the quadratic expression
$$3x^2 + 20x - 7$$
has the value zero when $x = -7$; hence, $(x + 7)$ is a factor.
By further trial and error, the complete factorisation is
$$(x + 7)(3x - 1).$$

(b) Cubic Expressions

The method just used for the factorisation of quadratic expressions may also be used for polynomials having powers of x higher than two. In particular, it may be used for cubic expressions whose standard form is
$$ax^3 + bx^2 + cx + d,$$
where a, b, c and d are constants.

EXAMPLES

1. Factorise the cubic expression
$$x^3 + 3x^2 - x - 3$$
assuming that there is at least one whole number value of x for which the expression has the value zero.
Solution
By trial and error, the cubic expression has the value zero when $x = 1$. Hence, by the Factor Theorem, $(x - 1)$ is a factor.
Thus,
$$x^3 + 3x^2 - x - 3 \equiv (x - 1)Q(x)$$

where $Q(x)$ is some quadratic expression to be found and, if possible, factorised further. In other words,

$$x^3 + 3x^2 - x - 3 \equiv (x - 1)(px^2 + qx + r)$$

for some constants p, q and r.

Usually, it is fairly easy to determine $Q(x)$ using trial and error by comparing, initially, the highest and lowest powers of x on both sides of the above identity; then comparing any intermediate powers as necessary. We obtain

$$x^3 + 3x^2 - x - 3 \equiv (x - 1)(x^2 + 4x + 3).$$

In this example, the quadratic expression does factorise even further giving

$$x^3 + 3x^2 - x - 3 \equiv (x - 1)(x + 1)(x + 3).$$

2. Factorise the cubic expression

$$x^3 + 4x^2 + 4x + 1.$$

Solution

By trial and error, we discover that the cubic expression has value zero when $x = -1$, and so $x + 1$ must be a factor.

Hence,

$$x^3 + 4x^2 + 4x + 1 \equiv (x + 1)Q(x),$$

where $Q(x)$ is a quadratic expression to be found. In other words,

$$x^3 + 4x^2 + 4x + 1 \equiv (x + 1)(px^2 + qx + r)$$

for some constants p, q and r.

Comparing the relevant coefficients, we obtain

$$x^3 + 4x^2 + 4x + 1 \equiv (x + 1)(x^2 + 3x + 1)$$

but, this time, the quadratic part of the answer will not conveniently factorise further.

1.8.3 CUBIC EQUATIONS

There is no convenient general method of solving a cubic equation; hence, the discussion here will be limited to those equations where at least one of the solutions is known to be a fairly small whole number, that solution being obtainable by trial and error.

We illustrate with examples.

EXAMPLES

1. Solve the cubic equation
$$x^3 + 3x^2 - x - 3 = 0.$$

   Solution
   By trial and error, one solution is $x = 1$ and so $(x - 1)$ must be a factor of the left hand side. In fact, we obtain the new form of the equation as
$$(x - 1)(x^2 + 4x + 3) = 0.$$
   That is,
$$(x - 1)(x + 1)(x + 3) = 0.$$
   Hence, the solutions are $x = 1$, $x = -1$ and $x = -3$.

2. Solve the cubic equation
$$2x^3 - 7x^2 + 5x + 54 = 0.$$

   Solution
   By trial and error, one solution is $x = -2$ and so $(x + 2)$ must be a factor of the left hand side. In fact we obtain the new form of the equation as
$$(x + 2)(2x^2 - 11x + 27) = 0.$$
   The quadratic factor will not conveniently factorise, but we can use the quadratic formula to determine the values of x, if any, which make it equal to zero.They are
$$x = \frac{11 \pm \sqrt{121 - 216}}{4}.$$
   The discriminant here is negative so that the only (real) solution to the cubic equation is $x = -2$.

1.8.4 LONG DIVISION OF POLYNOMIALS

(a) Exact Division

Having used the Factor Theorem to find a linear factor of a polynomial expression, there will always be a remaining factor, $Q(x)$, which is some other polynomial whose degree is one lower than that of the original; but the determination of $Q(x)$ by trial and error is not the only method of doing so. An alternative method is to use the technique known as **"long division of polynomials"**; and the working is illustrated by the following examples:

EXAMPLES

1. Factorise the cubic expression
$$x^3 + 3x^2 - x - 3.$$

   Solution

   By trial and error, the cubic expression has the value zero when $x = 1$ so that $(x - 1)$ is a factor.

   Dividing the given cubic expression (called the **"dividend"**) by $(x - 1)$ (called the **"divisor"**), we have the following scheme:

$$\begin{array}{r@{}l}
 & x^2 + 4x + 3 \\ \cline{2-2}
x - 1) & x^3 + 3x^2 - x - 3 \\
 & \underline{x^3 - x^2} \\
 & 4x^2 - x - 3 \\
 & \underline{4x^2 - 4x} \\
 & 3x - 3 \\
 & \underline{3x - 3} \\
 & 0
\end{array}$$

   Note:
   At each stage, using positive powers of x or constants only, we examine what the highest power of x in the divisor would have to be multiplied by in order to give the highest power of x in the dividend. The results are written in the top line as the **"quotient"** so that, on multiplying down, we can subtract to find each **"remainder"**.
   The process stops when, to continue would need negative powers of x; the final remainder will be zero when the divisor is a factor of the original expression.

   We conclude here that
$$x^3 + 3x^2 - x - 3 \equiv (x - 1)(x^2 + 4x + 3).$$

   Further factorisation leads to the complete result,
$$x^3 + 3x^2 - x - 3 \equiv (x - 1)(x + 1)(x + 3).$$

2. Solve, completely, the cubic equation
$$x^3 + 4x^2 + 4x + 1 = 0.$$

   Solution

   By trial and error, one solution is $x = -1$ so that $(x + 1)$ is a factor of the left hand side.

Dividing the left hand side of the equation by $(x+1)$, we have the following scheme:

$$\begin{array}{r l} & x^2 + 3x + 1 \\ x+1) & \overline{x^3 + 4x^2 + 4x + 1} \\ & \underline{x^3 + x^2} \\ & \quad 3x^2 + 4x + 1 \\ & \quad \underline{3x^2 + 3x} \\ & \qquad\quad x + 1 \\ & \qquad\quad \underline{x + 1} \\ & \qquad\qquad 0 \end{array}$$

Hence, the equation becomes

$$(x+1)(x^2+3x+1) = 0,$$

giving $x = -1$ and $x = \frac{-3\pm\sqrt{9-4}}{2} \simeq -0.382 \quad \text{or} \quad -2.618$

(b) Non-exact Division

The technique for long division of polynomials may be used to divide any polynomial by another polynomial of lower or equal degree, even when this second polynomial is not a factor of the first.

The chief difference in method is that the remainder will not be zero; but otherwise we proceed as before.

EXAMPLES

1. Divide the polynomial $6x+5$ by the polynomial $3x-1$.

   Solution

$$\begin{array}{r l} & \quad 2 \\ 3x-1) & \overline{6x + 5} \\ & \underline{6x - 2} \\ & \quad 7 \end{array}$$

Hence,

$$\frac{6x+5}{3x-1} \equiv 2 + \frac{7}{3x-1}.$$

2. Divide $3x^2 + 2x$ by $x + 1$.

   Solution

$$\begin{array}{r}
3x - 1 \\
x + 1 \overline{\smash{)}\, 3x^2 + 2x} \\
\underline{3x^2 + 3x} \\
-x \\
\underline{-x - 1} \\
1
\end{array}$$

   Hence,

$$\frac{3x^2 + 2x}{x + 1} \equiv 3x - 1 + \frac{1}{x + 1}.$$

3. Divide $x^4 + 2x^3 - 2x^2 + 4x - 1$ by $x^2 + 2x - 3$.

   Solution

$$\begin{array}{r}
x^2 \qquad\quad + 1 \\
x^2 + 2x - 3 \overline{\smash{)}\, x^4 + 2x^3 - 2x^2 + 4x - 1} \\
\underline{x^4 + 2x^3 - 3x^2} \qquad\qquad \\
x^2 + 4x - 1 \\
\underline{x^2 + 2x - 3} \\
2x + 2
\end{array}$$

   Hence,

$$\frac{x^4 + 2x^3 - 2x^2 + 4x - 1}{x^2 + 2x - 3} \equiv x^2 + 1 + \frac{2x + 2}{x^2 + 2x - 3}.$$

1.8.5 EXERCISES

1. Use the Factor Theorem to factorise the following quadratic expressions: assuming that at least one root is an integer:
   (a) $2x^2 + 7x + 3$;
   (b) $1 - 2x - 3x^2$.
2. Factorise the following cubic polynomials assuming that at least one root is an integer:
   (a) $x^3 + 2x^2 - x - 2$;
   (b) $x^3 - x^2 - 4$;
   (c) $x^3 - 4x^2 + 4x - 3$.

3. Solve completely the following cubic equations assuming that at least one solution is an integer:

   (a) $x^3 + 6x^2 + 11x + 6 = 0$;

   (b) $x^3 + 2x^2 - 31x + 28 = 0$.

4. (a) Divide $2x^3 - 11x^2 + 18x - 8$ by $x - 2$;

   (b) Divide $3x^3 + 12x^2 + 13x + 4$ by $x + 1$.

5. (a) Divide $x^2 - 2x + 1$ by $x^2 + 3x - 2$, expressing your answer in the form of a constant plus a ratio of two polynomials.

   (b) Divide x^5 by $x^2 - 2x - 8$, expressing your answer in the form of a polynomial plus a ratio of two other polynomials.

1.8.6 ANSWERS TO EXERCISES

1. (a) $(2x + 1)(x + 3)$; (b) $(x + 1)(1 - 3x)$.
2. (a) $(x - 1)(x + 1)(x + 2)$; (b) $(x - 2)(x^2 + x + 2)$; (c) $(x - 3)(x^2 - x + 1)$.
3. (a) $x = -1$, $x = -2$, $x = -3$; (b) $x = 1$, $x = 4$, $x = -7$.
4. (a) $2x^2 - 7x + 4$; (b) $3x^2 + 9x + 4$.
5. (a)

$$1 + \frac{3 - 5x}{x^2 + 3x - 2};$$

   (b)

$$x^3 + 2x^2 + 12x + 40 + \frac{176x + 320}{x^2 - 2x - 8}.$$

"JUST THE MATHS"

UNIT NUMBER

1.9

ALGEBRA 9
(The theory of partial fractions)

by

A.J.Hobson

UNIT 1.9 - ALGEBRA 9 - THE THEORY OF PARTIAL FRACTIONS

1.9.1 INTRODUCTION

The theory of partial fractions applies chiefly to the ratio of two polynomials in which the degree of the numerator is strictly less than that of the denominator. Such a ratio is called a **"proper rational function"**.

For a rational function which is not proper, it is necessary first to use long division of polynomials in order to express it as the sum of a polynomial and a proper rational function.

RESULT

A proper rational function whose denominator has been factorised into its irreducible factors can be expressed as a sum of proper rational functions, called **"partial fractions"**; the denominators of the partial fractions are the irreducible factors of the denominator in the original rational function.

ILLUSTRATION

From previous work on fractions, it can be verified that

$$\frac{1}{2x+3}+\frac{3}{x-1}\equiv\frac{7x+8}{(2x+3)(x-1)}$$

and the expression on the left hand side may be interpretted as the decomposition into partial fractions of the expression on the right hand side.

1.9.2 STANDARD TYPES OF PARTIAL FRACTION PROBLEM

(a) Denominator of the given rational function has all linear factors.

EXAMPLE

Express the rational function

$$\frac{7x+8}{(2x+3)(x-1)}$$

in partial fractions.

Solution

There will be two partial fractions each of whose numerator must be of lower degree than 1; i.e. it must be a **constant**

We write

$$\frac{7x+8}{(2x+3)(x-1)}\equiv\frac{A}{2x+3}+\frac{B}{x-1}.$$

Multiplying throughout by $(2x + 3)(x - 1)$, we obtain

$$7x + 8 \equiv A(x - 1) + B(2x + 3).$$

In order to determine A and B, any two suitable values of x may be substituted on both sides; and the most obvious values in this case are $x = 1$ and $x = -\frac{3}{2}$.

It may, however, be argued that these two values of x must be disallowed since they cause denominators in the first identity above to take the value zero.

Nevertheless, we shall use these values in the second identity above since the arithmetic involved is negligibly different from taking values infinitesimally close to $x = 1$ and $x = -\frac{3}{2}$.

Substituting $x = 1$ gives

$$7 + 8 = B(2 + 3).$$

Hence,

$$B = \frac{7 + 8}{2 + 3} = \frac{15}{5} = 3.$$

Substituting $x = -\frac{3}{2}$ gives

$$7 \times -\frac{3}{2} + 8 = A(-\frac{3}{2} - 1).$$

Hence,

$$A = \frac{7 \times -\frac{3}{2} + 8}{-\frac{3}{2} - 1} = \frac{-\frac{5}{2}}{-\frac{5}{2}} = 1.$$

We conclude that

$$\frac{7x + 8}{(2x + 3)(x - 1)} = \frac{1}{2x + 3} + \frac{3}{x - 1}.$$

The "Cover-up" Rule
A useful time-saver when the factors in the denominator of the given rational function are linear is to use the following routine which is equivalent to the method described above:

To obtain the constant numerator of the partial fraction for a particular linear factor, $ax + b$, in the original denominator, cover up $ax + b$ in the original rational function and then substitute $x = -\frac{b}{a}$ into what remains.

ILLUSTRATION
In the above example, we may simply cover up $x - 1$, then substitute $x = 1$ into the fraction

$$\frac{7x + 8}{2x + 3}.$$

Then we may cover up $2x+3$ and substitute $x=-\frac{3}{2}$ into the fraction

$$\frac{7x+8}{x-1}.$$

Note:
We shall see later how the cover-up rule can also be brought into effective use when not all of the factors in the denominator of the given rational function are linear.

(b) Denominator of the given rational function contains one linear and one quadratic factor

EXAMPLE

Express the rational function

$$\frac{3x^2+9}{(x-5)(x^2+2x+7)}$$

in partial fractions.

Solution

We should observe firstly that the quadratic factor will not reduce conveniently into two linear factors. If it did, the method would be as in the previous paragraph. Hence we may write

$$\frac{3x^2+9}{(x-5)(x^2+2x+7)} \equiv \frac{A}{x-5} + \frac{Bx+C}{x^2+2x+7},$$

noticing that the second partial fraction may contain an x term in its numerator, yet still be a proper rational function.

Multiplying throughout by $(x-5)(x^2+2x+7)$, we obtain

$$3x^2+9 \equiv A(x^2+2x+7)+(Bx+C)(x-5).$$

A convenient value of x to substitute on both sides is $x=5$ which gives

$$3\times 5^2+9 = A(5^2+2\times 5+7).$$

That is, $84=42A$ or $A=2$.

No other convenient values of x may be substituted; but two polynomial expressions can be identical only if their corresponding coefficients are the same in value. We therefore equate suitable coefficients to find B and C; usually, the coefficients of the highest and lowest powers of x.

Equating coefficients of x^2, $3 = A + B$ and hence $B = 1$.

Equating constant terms (the coefficients of x^0), $9 = 7A - 5C = 14 - 5C$ and hence $C = 1$.

The result is therefore

$$\frac{3x^2+9}{(x-5)(x^2+2x+7)} \equiv \frac{2}{x-5} + \frac{x+1}{x^2+2x+7}.$$

Observations
It is easily verified that the value of A may be calculated by means of the cover-up rule, as in paragraph **(a)**; and, having found A, the values of B and C could be found by cross multiplying the numerators and denominators in the expression

$$\frac{2}{x-5} + \frac{?x+?}{x^2+2x+7}$$

in order to arrive at the numerator of the original rational function. This process essentially compares the coefficients of x^2 and x^0 as before.

(c) Denominator of the given rational function contains a repeated linear factor

In general, examples of this kind will not be more complicated than for a rational function with one repeated linear factor together with either a non-repeated linear factor or a quadratic factor.

EXAMPLE

Express the rational function

$$\frac{9}{(x+1)^2(x-2)}$$

in partial fractions.

Solution

First we observe that, from paragraph **(b)**, the partial fraction corresponding to the repeated linear factor would be of the form

$$\frac{Ax+B}{(x+1)^2};$$

but this may be written

$$\frac{A(x+1)+B-A}{(x+1)^2} \equiv \frac{A}{x+1} + \frac{B-A}{(x+1)^2} \equiv \frac{A}{x+1} + \frac{C}{(x+1)^2}.$$

Thus, a better form of statement for the problem as a whole, is

$$\frac{9}{(x+1)^2(x-2)} \equiv \frac{A}{x+1} + \frac{C}{(x+1)^2} + \frac{D}{x-2}.$$

Eliminating fractions, we obtain

$$9 \equiv A(x+1)(x-2) + C(x-2) + D(x+1)^2.$$

Putting $x = -1$ gives $9 = -3C$ so that $C = -3$.

Putting $x = 2$ gives $9 = 9D$ so that $D = 1$.

Equating coefficients of x^2 gives $0 = A + D$ so that $A = -1$.

Therefore,

$$\frac{9}{(x+1)^2(x-2)} \equiv -\frac{1}{x+1} - \frac{3}{(x+1)^2} + \frac{1}{x-2}.$$

Notes:
(i) Similar partial fractions may be developed for higher repeated powers so that, for a repeated linear factor of power of n, there will be n corresponding partial fractions, each with a constant numerator. The labels for these numerators in future will be taken as A, B, C, etc. in sequence.

(ii) We observe that the numerator above the repeated factor itself (D in this case) could actually have been obtained by the cover-up rule; covering up $(x+1)^2$ in the original rational function, then substituting $x = -1$ into the rest.

(d) Keily's Method

A useful method for repeated linear factors is to use these factors one at a time, keeping the rest outside the expression as a factor.

EXAMPLE

Express the rational function

$$\frac{9}{(x+1)^2(x-2)} \equiv \frac{1}{x+1}\left[\frac{9}{(x+1)(x-2)}\right]$$

in partial fractions.

Solution

Using the cover-up rule inside the square brackets,

$$\frac{9}{(x+1)^2(x-2)} \equiv \frac{1}{x+1}\left[\frac{-3}{x+1}+\frac{3}{x-2}\right]$$

$$\equiv -\frac{3}{(x+1)^2}+\frac{3}{(x+1)(x-2)};$$

and, again by cover-up rule,

$$\equiv -\frac{3}{(x+1)^2}-\frac{1}{x+1}+\frac{1}{x-2}$$

as before.

Warning
Care must be taken with Keily's method when, even though the original rational function is proper, the resulting expression inside the square brackets is improper. This would have occurred, for instance, if the problem given had been

$$\frac{9x^2}{(x+1)^2(x-2)},$$

leading to

$$\frac{1}{x+1}\left[\frac{9x^2}{(x+1)(x-2)}\right].$$

In this case, long division would have to be used inside the square brackets before proceeding with Keily's method.

For such examples, it is probably better to use the method of paragraph **(c)**.

1.9.3 EXERCISES

Express the following rational functions in partial fractions:

1.

$$\frac{3x+5}{(x+1)(x+2)}.$$

2.

$$\frac{17x+11}{(x+1)(x-2)(x+3)}.$$

3.

$$\frac{3x^2-8}{(x-1)(x^2+x-7)}.$$

4.

$$\frac{2x+1}{(x+2)^2(x-3)}.$$

5.

$$\frac{9+11x-x^2}{(x+1)^2(x+2)}.$$

6.

$$\frac{x^5}{(x+2)(x-4)}.$$

1.9.4 ANSWERS TO EXERCISES

1.

$$\frac{2}{x+1}+\frac{1}{x+2}.$$

2.

$$\frac{1}{x+1}+\frac{3}{x-2}-\frac{4}{x+3}.$$

3.

$$\frac{1}{x-1}+\frac{2x+1}{x^2+x-7}.$$

4.

$$\frac{3}{5(x+2)^2}-\frac{7}{25(x+2)}+\frac{7}{25(x-3)}.$$

5.

$$-\frac{3}{(x+1)^2}+\frac{16}{x+1}-\frac{17}{x+2}.$$

6.

$$x^3+2x^2+12x+40+\frac{16}{3(x+2)}+\frac{512}{3(x-4)}.$$

"JUST THE MATHS"

UNIT NUMBER

1.10

ALGEBRA 10
(Inequalities 1)

by

A.J.Hobson

UNIT 1.10 - ALGEBRA 10 - INEQUALITIES 1.

1.10.1 INTRODUCTION

If the symbols a and b denote numerical quantities, then the statement

$$a < b$$

is used to mean "a is less than b" while the statement

$$b > a$$

is used to mean "b is greater than a"

These are called **"strict inequalities"** because there is no allowance for the possibilty that a and b might be equal to each other. For example, if a is the number of days in a particular month and b is the number of hours in that month, then $b > a$.

Some inequalities do allow the possibility of a and b being equal to each other and are called **"weak inequalities"** written in one of the forms

$$a \leq b \quad b \leq a \quad a \geq b \quad b \geq a$$

For example, if a is the number of students who enrolled for a particular module in a university and b is the number of students who eventually passed that module, then $a \geq b$.

1.10.2 ALGEBRAIC RULES FOR INEQUALITIES

Given two different numbers, one of them must be strictly less than the other. Suppose a is the smaller of the two and b the larger; i.e.

$$a < b$$

Then
1. $a + c < b + c$ for any number c.

2. $ac < bc$ when c is positive but $ac > bc$ when c is negative.

3. $\frac{1}{a} > \frac{1}{b}$ provided a and b are **both positive.**

Note:
The only other situation in 3. which is consistent with $a < b$ will occur if a is negative and b is positive. In this case, $\frac{1}{a} < \frac{1}{b}$ because a negative number is always less than a positive number.

EXAMPLES

1. Simplify the inequality
$$2x + 3y > 5x - y + 7.$$

 Solution

 We simply deal with this in the same way as we would deal with an equation by adding appropriate quantities to both sides or subtracting appropriate quantities from both sides. We obtain
$$-3x + 4y > 7 \quad \text{or} \quad 3x - 4y + 7 < 0.$$

2. Solve the inequality
$$\frac{1}{x-1} < 2,$$
 assuming that $x \neq 1$.

 Solution

 Here we must be careful in case $x-1$ is negative; the argument is therefore in two parts:
 (a) If $x > 1$, i.e. $x - 1$ is positive, then the inequality can be rewritten as
$$1 < 2(x-1) \quad \text{or} \quad x - 1 > \frac{1}{2}.$$
 Hence,
$$x > \frac{3}{2}.$$
 (b) If $x < 1$, i.e. $x - 1$ is negative, then the inequality is automatically true since a negative number is bound to be less than a positive number.

 Conclusion:
 The given inequality is satisfied when $x < 1$ and when $x > \frac{3}{2}$.

3. Solve the inequality
$$2x - 7 \leq 3.$$

 Solution

 Adding 7 to both sides, then dividing both sides by 2 gives
$$x \leq 5.$$

4. Solve the inequality
$$\frac{x-1}{x-6} \geq 0.$$

Solution

We first observe that the fraction on the left of the inequality can equal zero only when $x = 1$.

Secondly, the only way in which a fraction can be positive is for both numerator and denominator to be positive or both numerator and denominator to be negative.

(a) Suppose $x - 1 > 0$ and $x - 6 > 0$; these two are covered by $x > 6$.

(b) Suppose $x - 1 < 0$ and $x - 6 < 0$; these two are covered by $x < 1$.

Note:
The value $x = 6$ is problematic because the given expression becomes infinite; in fact, as x passes through 6 from values below it to values above it, there is a sudden change from $-\infty$ to $+\infty$.

Conclusion:
The inequality is satisfied when either $x > 6$ or $x \leq 1$.

1.10.3 INTERVALS

In scientific calculations, a variable quantity x may be restricted to a certain range of values called an **"interval"** which may extend to ∞ or $-\infty$; but, in many cases, such intervals have an upper and a lower **"bound"**. The standard types of interval are as follows:

(a) $a < x < b$ denotes an **"open interval"** of all the values of x between a and b but excluding a and b themselves. The symbol (a, b) is also used to mean the same thing. For example, if x is a purely decimal quantity, it must lie in the open interval

$$-1 < x < 1.$$

(b) $a \leq x \leq b$ denotes a **"closed interval"** of all the values of x from a to b inclusive. The symbol $[a, b]$ is also used to mean the same thing. For example, the expression $\sqrt{1 - x^2}$ has real values only when

$$-1 \leq x \leq 1.$$

Note:
It is possible to encounter intervals which are closed at one end but open at the other; they may be called either **"half open"** or **"half closed"**. For example

$$a < x \leq b \quad \text{or} \quad a \leq x < b,$$

which can also be denoted respectively by $(a, b]$ and $[a, b)$.

(c) Intervals of the types

$$x > a \quad x \geq a \quad x < a \quad x \leq a$$

are called **"infinite intervals"**.

1.10.4 EXERCISES

1. Simplify the following inequalities:

   (a)

   $$x + y \leq 2x + y + 1;$$

   (b)

   $$2a - b > 1 + a - 2b - c.$$

2. Solve the following inequalities to find the range of values of x.

   (a)

   $$x + 3 < 6;$$

   (b)

   $$-2x \geq 10;$$

   (c)

   $$\frac{2}{x} > 18;$$

   (d)

   $$\frac{x+3}{2x-1} \leq 0.$$

3. Classify the following intervals as open, closed, or half-open/half-closed:

   (a)

   $$(5, 8);$$

   (b)

   $$(-3, -2);$$

(c)

$$[2, 4);$$

(d)

$$[8, 23];$$

(e)

$$(-\infty, \infty);$$

(f)

$$(0, \infty);$$

(g)

$$[0, \infty).$$

1.10.5 ANSWERS TO EXERCISES

1. (a) $x \geq -1$;
   (b) $a + b + c > 1$.
2. (a) $x < 3$;
   (b) $x \leq -5$;
   (c) $x < \frac{1}{9}$ since $x > 0$;
   (d) $-3 \leq x < \frac{1}{2}$.
3. (a) open;
   (b) open;
   (c) half-open/half-closed;
   (d) closed;
   (e) open;
   (f) open;
   (g) half-open/half-closed.

"JUST THE MATHS"

UNIT NUMBER

1.11

ALGEBRA 11
(Inequalities 2)

by

A.J.Hobson

UNIT 1.11 - ALGEBRA 11 - INEQUALITIES 2.

1.11.1 RECAP ON MODULUS, ABSOLUTE VALUE OR NUMERICAL VALUE

As seen in Unit 1.1, the Modulus of a numerical quantity ignores any negative signs if there are any. For example, the modulus of -3 is 3, but the modulus of 3 is also 3.

The modulus of an unspecified numerical quantity x is denoted by the symbol

$$\mid x \mid$$

and is defined by the two statements:

$$\mid x \mid = x \quad \text{if} \quad x \geq 0;$$

$$\mid x \mid = -x \quad \text{if} \quad x \leq 0.$$

Notes:
(i) An alternative, but less convenient formula for the modulus of x is

$$\mid x \mid = +\sqrt{x^2}.$$

(ii) It is possible to show that, for any two numbers a and b,

$$\mid a + b \mid \leq \mid a \mid + \mid b \mid .$$

This is called the **"triangle inequality"** and can be linked to the fact that the length of any side of a triangle is never greater than the sum of the lengths of the other two sides.

The proof is a little involved since it is necessary to consider all possible cases of a and b being positive, negative or zero together with a consideration of their relative sizes. It will not be included here.

1.11.2 INTERVAL INEQUALITIES

(a) Using the Modulus notation

In this section, we investigate the meaning of the inequality

$$\mid x - a \mid < k,$$

where a is any number and k is a positive number.