ILLUSTRATIONS IN COLOUR

ILLUSTRATIONS

IN BLACK AND WHITE

Editorial

THIS is our fifteenth issue. It is the largest yet to appear because it contains as a supplement a complete subject and author index covering everything from number one in June 1946 to number fifteen here and now. It makes a convenient moment to pause and survey our work.

Most people imagine that an editor's life consists in sitting in an armchair wielding a blue pencil, while every post brings a flood of new contributions to be sorted and rejected – a soft, easy, parasite existence, just reading whatever the seas wash up. For most periodicals, and *Science News* is no exception, nothing could be further from the truth. The editors of *Horizon* and *Penguin New Writing* have publicly remarked on the hard work they do wringing contributions out of authors, and other editors I have talked with say the same thing: if they had to rely on the free-lance articles which come in they could close down tomorrow.

There are many reasons for this, of course. People wait to be invited to write because they are too proud or too modest, too busy or too lazy to offer an article of their own accord. Some of the contributions which do come in to us are so badly written, so muddled and full of jargon, that an editorial attempt to reduce the chaos and translate into English is just not worth while. The inability of most British scientists to write even a proper scientific paper, let alone a popular article, is now causing such general concern that a number of big firms and scientific societies are appointing advisers in English literature to help their inarticulate colleagues to express themselves intelligibly and with precision. It will take a long time for this movement to produce much effect, and it is a question whether the regular writing of English essays should not become a part of the university training of every science undergraduate.

But when all this is said there is one fundamental reason why free-lance articles so rarely see print. Every magazine is created

Plate 1. Red *Daphnia*, with much haemoglobin in its blood. × 18.
(See the article on 'The Blood of Water Fleas' by
Professor H. Munro Fox, F.R.S.)

(Plates 1, 2 and 3 reproduced by permission of *Endeavour*)

SCIENCE NEWS

EDITED BY J. L. CRAMMER

15

PENGUIN BOOKS

HARMONDSWORTH · MIDDLESEX

RE. 1-2-0
THE PRESIDENCY BOOK SUPPLIES
TRIPLICANE, MADRAS 5.

These facts caused Thomas Young, who has since been called 'the father of physiological optics', to propose, a century and a half ago, the three-colour theory of vision which bears his name. This theory may be stated in various ways, one of which is as follows: there cannot be an infinite number of different kinds of nerve fibre connecting the retina* with the brain, so let us suppose the number of kinds to be three, and that the sensations of redness are conveyed by one variety, the sensations of greenness by another variety, and the sensations of blueness by a third variety. In addition there must be three kinds of retinal light receptor for coloured light rays, namely those for red rays, those for green rays and those for blue rays. The incidence of red light on the retina in sufficient amount will act as a stimulus for the red receptors and will cause nerve impulses to travel up the optic nerve fibres to the brain, so that a red sensation is perceived in consequence. The incidence of green rays or of blue rays will similarly stimulate the green or the blue receptors and will thus arouse green or blue sensations. 'Let us suppose further,' said Young, 'that yellow light stimulates both the red and the green receptors at the same time, and that when this happens a red sensation and a green sensation are aroused, producing as a combined effect a sensation of yellowness. If then green rays and blue rays together cause a blue-green sensation, blue rays and red rays together cause a purple, mauve or violet sensation, and all three together a white sensation, we have a mechanism which will account for many of the facts of colour vision.'

This theory did account in a satisfactory manner for all the facts known in Thomas Young's time, and it has accounted for much which has been discovered since, so that it would be true to say that this theory has played an extremely important part in the science of colour vision during the past century; for not only has it offered a feasible explanation of this important function of the eye, but it had led to such

*The sensitive surface inside the eye, corresponding to the film in a camera.

valuable practical developments as colour photography and Technicolor.

Recently, however, doubts concerning the three-colour theory have arisen, and the idea has been gaining ground that things are not quite as simple as they appeared at first. How have these doubts arisen? As far as the writer is concerned they were brought to his notice in the following way.

One of the most puzzling features as far as the lens system of the human eye is concerned has been two apparently contradictory sets of facts. On the one hand there has been the knowledge that the lens system of the eye is not corrected for chromatic aberration. This means that if the retina is situated at the focusing point for the yellow rays then the blue and violet rays will come to a focus in front of it, and the red and orange rays will come to a focus behind it. The image of a point source will then consist of a bright yellow centre surrounded by a reddish blue halo. On the other hand we have the evidence of our eyes when they are being used for examining everyday objects, that when we look at a distant point source of light we do not see a yellow point of light, neither do we see surrounding it the reddish blue halo. How is this contradiction to be explained?

Let us begin by examining the evidence a little more closely. Measurements of the amount of chromatic aberration have been made by many workers. They are all agreed that it is present and they are even agreed as to its amount, except for very small differences of detail. There are also many physiological experiments which are designed to show that this aberration is present. Two may be mentioned in this place.

(*a*) On viewing the filaments of an electric lamp through a piece of blue or purple glass, which transmits at the same time some red rays and some blue rays, but cuts out the rays from the middle of the spectrum, it is found either that the red rays are in focus, and that the blue rays form a halo round them, or that the blue rays are in focus and that the red rays form a halo round them, clearly showing that it is

impossible to focus both sets of rays simultaneously because of the differences in the position of their focusing points. Occasionally a person is found who does not focus either the red rays or the blue rays sharply, in which case both of them form blurs. This condition is due to the fact that the blue rays have come to a focus in front of the retina and are now separating again, whereas the red rays have not yet come to a focus.

(*b*) There is a further point to be mentioned, namely that a careful examination of the refractive media of the eye (the cornea and the lens) fails altogether to disclose the presence of a method of correcting this aberration similar to that which is employed in optical instruments. There seems, therefore, to be overwhelming evidence that chromatic aberration is present.

Our next step is to examine the effects which chromatic aberration would have, and this may be done by employing a telescope or a microscope in which the usual objective which has been corrected for colour is replaced by one which does not possess this correction. It is then found that the images presented to the observer's eye consist of coloured spots and haloes and fringes. Thus a distant point of white light appears to be a bright yellow spot surrounded by a reddish blue halo, and a large black object on a white ground appears to have at its margin two narrow fringes, one yellow and the other violet, which surround the black object on all sides. If a grating consisting of black and white bars be examined instead, then the white bars are found to be portrayed in yellow and the black bars in blue or violet. It is clear that in all these cases chromatic aberration produces colour effects of an unmistakable kind which cannot be tolerated in a well-designed optical instrument.

Now, as mentioned above, when the eye is used for normal vision no indication of these coloured fringes can be observed. Moreover the use of monochromatic light, for example sodium light, does not improve the acuity of the eye to more than a very small extent, neither does the correction of the eye by means of an external arrangement of lenses through which the light passes on its way into the eye noticeably improve definition. The conclusion to which we are forced is, therefore, that *in some way* the chromatic aberration of the eye is corrected or eliminated, so that the coloured fringes produced by the lens system of the eye are removed again. The problem is, therefore, to ascertain how this is done.

Many years ago the writer was engaged in examining this problem when he made a curious observation. It was found that short lengths of white thread and short lengths of yellow thread, when placed on a black background, could readily be distinguished from one another when they were close to the eye, but that they could no longer be distinguished if the distance between them and the observer's eye was increased by a suitable amount. It appeared from this experiment that there was some mechanism by means of which the yellowness of the threads had been eliminated, causing them to resemble the white threads.

A similar experiment was performed using blue threads and black threads. When these were placed on a white background it was found that they were readily distinguished when they were close to the observer, but that they could no longer be distinguished if they were placed at a sufficient distance from him. Now the essential difference between yellow and white is that the former reflects all colours of the spectrum except blue and violet, whereas white reflects all colours of the spectrum, including blue and violet. If, then, the sensations of blue and violet were in some way put out of action it would cause yellow and white to resemble one another, as experiment shows to be the case.

With regard to the blue and black threads, in this case also the similarity would be accounted for if in some manner the blue receptors were put out of action. The writer, therefore, put forward the suggestion that the confusion of yellow with white and the confusion of blue with black might be accounted for on Young's theory by supposing that some

change took place in the sensitivity of the blue-receiving mechanism of the eye. Now if this were the only change a moment's thought will show that, while accounting for the change undergone by blue, the hypothesis does not really account completely for the change undergone by yellow, because if blue reception be interfered with, white should become yellow in appearance. Now the writer's experiments showed with the greatest clearness that this did not happen, but that it was yellow which became white in appearance. A somewhat more elaborate hypothesis was therefore advanced, namely that there are two changes in the visual mechanism. The first is the reduction in activity of the blue receptor mechanism, as previously described, but a second change takes place at the same time, namely the introduction of an inter-connecting link between the green receptors and the brain centre for the blue sensation. With this new arrangement both yellow and white will stimulate the red receptors and the green receptors, and will thus in consequence of the cross link stimulate all three brain centres for red, green and blue, the result being that both will appear white, as is found by experiment to be the case.

This attractive hypothesis appeared to account for the problem of the chromatic aberration of the eye, as the following examples will show. Suppose the observer to be looking at a distant point source of white light. Chromatic aberration will cause the retinal image of this object to appear to consist of a bright yellow spot surrounded by a blue violet halo, but since the angle which these subtend at the eye is small, the visual mechanism is modified as described above, and the blue receptors cease to function and the green receptors cause nerve impulses to travel not only to the brain centre for the perception of green but to that for blue as well, the result being that the bright yellow centre is perceived as white and the blue violet halo appears to be replaced by black, and in consequence the observer sees the original object in a corrected manner.

Let us consider as a second example a grating consisting

of alternate black and white bars. Then the chromatic aberration of the eye causes the retinal image of this to consist of yellow bars alternating with blue violet ones, but yellow is replaced by white and blue violet by black, so that we have gone back to what we started with, and the observer sees what in fact is present, namely a grating consisting of black and white bars. By taking other examples the reader will find that in every case the chromatic aberration of the eye has been corrected.

This hypothesis based on Young's theory had unfortunately to be abandoned because of the discoveries made by other workers, Sloan, Wright, and Thomson. Sloan found that when colourless objects were examined in dim lights the yellow part of the spectrum suffered a marked reduction of brightness. Wright discovered that a similar reduction of brightness in the yellow part of the spectrum occurred when coloured objects were examined which subtended small angles at the eye. Lastly, Thomson has shown that not only is there a reduction of brightness in the yellow, but also one in the blue as well. Unfortunately these reductions of brightness cannot be accounted for on the hypothesis described above, the consequence being that an alternative hypothesis has to be considered.

Now it is clear that on Young's theory a reduction in brightness of the red part of the spectrum would also involve the orange and yellow parts in a reduction of brightness, and that similarly a reduction of the green would at the same time involve the yellow-green and yellow in a similar change; but what the three-colour theory cannot account for is the reduction of brightness of yellow alone, unaccompanied by a reduction in brightness either of red or of green.

Turning now to the short-wave part of the spectrum and examining the situation there, Thomson found the reduction of brightness to be limited to part of the spectrum having a wavelength* of about 0.45μ–0.47μ, and that parts of the spectrum having longer and shorter wavelengths than this were hardly affected at all. Now here again our observations

*For an explanation of *Wavelength* and the meaning of μ see *Glossary*.

do not fit in at all well with Young's theory, because if in fact the mechanism for blue were reduced in sensitivity there would be a reduction in brightness of the whole of the short wave part of the spectrum, and not just a reduction of a limited part only. It is difficult, therefore, to devise an explanation on the three-colour theory either for the changes which occur in the yellow or for the changes which occur in the blue, and in order to account for them adequately it is necessary to suppose that there are present in the human retina more kinds of receptor than three.

Five types of receptor would suffice for the purpose of correcting the chromatic aberration of the eye, namely red, yellow, green, blue and violet; the yellow and the blue being affected by a reduction of visual angle, whereas the receptors for the red, the green and the violet are not affected. Thus at ordinary visual angles all five types of receptors are in operation, but at small visual angles three types only, thus eliminating the yellow and the blue which would largely be responsible for the perception of the colour fringes which are produced by chromatic aberration.

II. The Significance of Vitamin A

It is a familiar fact that non-colour sensitive photographic plates and films darken, either directly or after development, after they have been exposed to ultra-violet light; that is, to rays which come from the part of the spectrum which is situated on the short wavelength side of the visible spectrum. When a photographic film is exposed in a spectroscope it is found that its sensitiveness extends into the blue and violet part of the visible spectrum. If a non-colour sensitive plate or film be bathed in a suitable dye before exposure it is caused to react to other parts of the visible spectrum, so that it becomes sensitive to visual rays because of the action of the dye. What happens is this; the dye absorbs the light and causes a chemical reaction to take place in the photographic emulsion. In the absence of the dye the light would not be absorbed, and in consequence exposing the plate to it would

produce no noticeable effect; thus dyes which absorb green rays cause the plate to be green sensitive; those which absorb orange and red rays cause it to be sensitive to these colours. When by the use of dyes it has been made sensitive to all visible rays it is known as a panchromatic plate or film. Examples similar to these show that there is a universal rule, namely that light rays cannot bring about a photo-chemical change unless they are absorbed. Thus when an attempt was made to prepare photographic plates which were sensitive to the infra-red or heat rays, dyes were employed which were strong absorbers for these rays.

The same principle must apply in the case of the eye; unless the receptors present in the retina either themselves absorb visible rays or are intimately associated with dyes which have the property of absorbing them, photochemical change cannot be brought about and the eye therefore remains blind. Thus in order that an animal may see, its retina must contain substances which have the property of absorbing light; that is pigments must be present. It has been suspected for many years that the coloured pigment 'visual purple', found in the retinas of such animals as frogs, is associated with twilight vision. This supposition has recently become a certainty, and it has been found that one of the chemical precursors of visual purple is vitamin A_1. Kühne found, when visual purple is acted on by light, that a new substance is formed which he called visual yellow. The latter has now been shown to be closely related to, if not identical with, the aldehyde of vitamin A_1, called retinene–1. When vitamin A_1 or retinene–1 are treated with strong acids such as sulphuric or phosphoric acid, Ball, Collins, Morton and Stubbs (1948) found that a number of coloured pigments are produced which have narrow well-marked absorption bands in the visible spectrum as shown in Table 1. Two facts about these derived pigments are important in view of what follows: (*a*) their limited absorption in a particular part of the spectrum and (*b*) the fact that they resemble visual purple in being bleached by ordinary white light.

Table 1.

Wave-length in μ	*Colour*	*Frog* Modulators*	*Vit. A* (H_2SO_4)	*Vit. A* (H_3PO_4)	*Retinene-1* (H_2SO_4)	*Retinene-1* (H_3PO_4)
0.66μ	Red	—	—	—	0.664μ	—
0.60μ	Orange	Found	0.62μ	0.60μ	—	—
0.58μ	Yellow	Found	0.59μ	—	—	0.59μ
0.54μ	Yellow Green	Found	—	—	0.56μ	0.55μ
0.5273μ	Green	Found	0.53μ	0.52μ	0.52μ	—
0.50μ	Blue Green	Not found	—	—	—	0.50μ
0.46μ	Blue	Found	0.465μ	0.44μ 0.48μ	0.44μ 0.46μ	0.47μ
0.44μ	Blue violet	Not found	—	—	—	—

*As found by Granit (1947), see section 3 below.

The significance of this work from the point of view of human vision lies in the fact that pigments other than visual purple can be prepared from vitamin A, which is found in the diet and is also found in various parts of the body, including the retina. If pigments such as those which have been prepared in the laboratory are present in the retina they will have the property of absorbing light, and the energy which they thus obtain would under favourable circumstances be available for acting as a stimulus to the sensory receptors. When these are stimulated in a suitable manner they have the property of sending out impulses to the brain and thus bring about the function of sight.

III. The Retinal Colour Receptors of Animals

It has been known for many years that, when nerve impulses travel along the fibres of a sensory or a motor nerve, changes of electrical potential take place, which are referred to as action currents.* When many nerve impulses are being transmitted the changes of potential are extremely

*See, for example, Dr Rushton's article in *Science News* 5.

rapid and appear on a photographic record as a series of spikes, rather like the teeth of a very fine saw. In order, therefore, to ascertain the events taking place in a nerve in the body, physiologists place suitable electrodes in contact with the nerve, so that these tap off a part of the change of electrical potential which is accompanying the transmission of the nerve impulses. In this way the messages transmitted by the nerve can be investigated under different conditions, and a knowledge can be gained of the kinds of messages which are being sent by the peripheral sense organs to the brain, or from the brain to the muscles and glands, in order to set them into activity.

In Granit's microelectrode method (1947) a fine platinum wire is placed in contact with one or more of the fibres of the optic nerve of an anaesthetised, or decerebrate animal, as they pass across the internal surface of the retina to emerge at the blind spot.* Monochromatic light of chosen wavelength, from a spectrometer, is now used in order to stimulate the retinal receptors. This light is decreased in intensity until only a few impulses pass up the nerve fibres on their way to the brain. When this is the case the changes of electrical potential which accompany the impulses cause potential changes in the platinum electrode, which is connected to a valve amplifier of high gain and a loud speaker. The latter emits a few clicks per minute when the light intensity has been correctly adjusted. Various wavelengths are investigated in turn until the complete response curve of the most sensitive receptor linked with the particular nerve fibre or fibres with which the microelectrode is in contact has been recorded. The electrode is now raised, and is placed in contact with another nerve fibre, whereupon the whole process is repeated.

The results obtained by Granit and his co-workers, using this technique, may be summarised as follows. There are three main types of response to a given light stimulus, 'on',

*For a brief account of the eye's structure, including blindspot fovea, etc., see *Glossary*.

'off' and 'on/off'. The 'on' response consists of a short volley of nerve impulses which accompanies the turning on of the light; the 'off' response consists similarly of a short volley which immediately follows the turning off of the light; the 'on/off' response consists of a brief volley of nerve impulses which occurs for a short time after either the onset of retinal illumination, or its cessation.

It was found that the receptor, or receptors, which are connected to any given optic nerve fibre respond to light rays within a certain narrow range of wavelengths. Thus one nerve might transmit impulses to the brain when the retina was stimulated by rays from the orange part of the spectrum, the maximum effect being produced by rays, for example, of 0.60μ. Another might be connected to a receptor with a narrow green response, and so on. The important facts about these responses were (*a*) that they were found to occur in many different parts of the spectrum; (*b*) that the response was limited to a narrow region of the spectrum; (*c*) that seven or eight different spectral regions were occupied by receptors, but that usually only a few of these were to be found in any one kind of animal (see Table 2); (*d*) that in addition to receptors with a narrow response in a particular region of the spectrum, which Granit called 'modulators', there are also receptors with a wide response, occupying almost the whole visible spectrum, which Granit called 'dominators'.

Table 2.

The modulators found in different animals (Granit).

Wave-length μ	*Cat*	*Guinea-pig*	*Rat*	*Snake*	*Frog*
0.60	Found	Found	Found	Found	Found
0.58	Found	—	—	—	Found
0.54	Found	—	—	—	Found
0.53	—	Found	—	Found	Found
0.50	—	Found	Found	—	—
0.48	—	—	—	—	—
0.46	Found	Found	—	—	Found
0.42	Found	—	—	—	—

The precise relationship between dominators and modulators is not yet known. It seems likely that the latter are responsible for performing an important role not only in connection with the sensation of colour, but also in connection with a sensation of brightness as well; thus modulators, which have a response in the red part of the spectrum, are not only responsible for the sensation of redness, but are also required in order to contribute the luminosity which is a property of this particular part of the spectrum. The dominators, on the contrary, which respond to all visible rays irrespective of their colour, must be non-colour-sensitive receptors. It seems likely that their main contribution, so far as sensation is concerned, is that of luminosity devoid of any colour association.

IV. The Retinal Direction Effects

In the field of human research two new facts have recently been discovered: (1) that the retina is more sensitive to light rays which are incident on it in a certain direction than it is to rays which reach it from other directions, and that the optimum direction varies in a complex manner with the wavelength of the incident light, (2) that various fixation points are employed for rays of varying wavelength, there being, according to the writer, as many as eight points, the approximate positions of which have been identified in the fovea.

In addition, two older fields of research have recently been re-explored: (3) the subjective colours which are seen when tiny pencils of white or coloured light are caused to explore the retina – this method, first used by Holmgren, has been revised and extended by the writer; (4) the colour properties of the fovea at small visual angles and at low light intensities.

There are really three different retinal direction effects: one on luminosity, one on hue and one on saturation; thus when a light ray which has passed through a peripheral part of the pupil is compared with one which has passed through a central part of this aperture, it is found to differ from it in

brightness, in colour and in the amount of white light which appears to be mixed with it. So far Stiles (1944) and his co-workers, to whom the discovery of these important direction effects is due, have investigated quantitatively only the effects on apparent brightness and on colour, and not those on saturation. The hue effect is shown diagramatically in Fig. 1 and the brightness effect in Fig. 2. It will be observed that

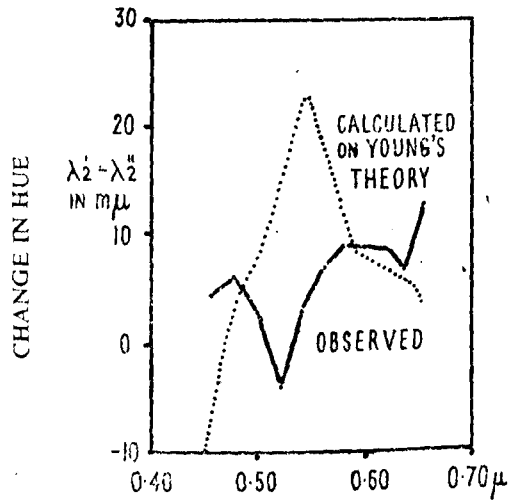


Fig. 1.

in both cases the various parts of the spectrum show marked individual differences so far as the magnitude of the effects is concerned. Attempts by Stiles to calculate the effects on the basis of the three-colour theory have led to failure, both in the case of brightness and in that of colour. Corresponding attempts by the writer to calculate the effects on the basis of the many types of receptor which are postulated by the polychromatic theory were found to be successful. These retinal effects constitute the first crucial test for the presence in the human retina of more types of colour receptor than the three postulated by Young's theory.

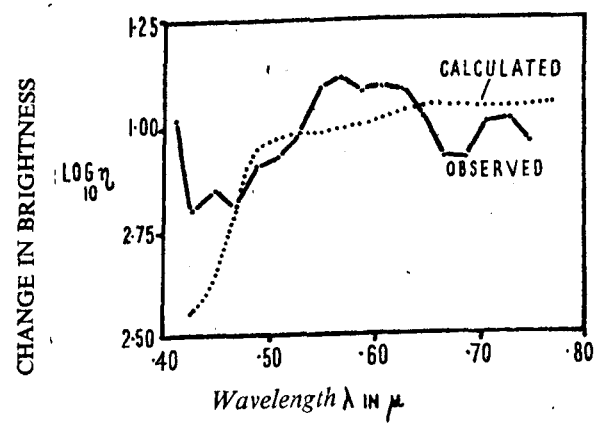


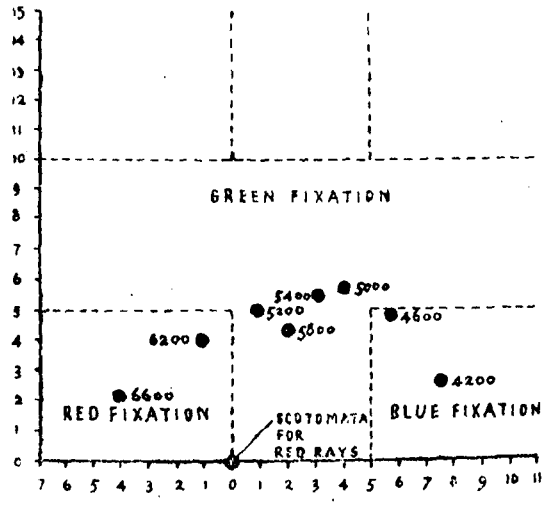
Fig. 2.

V. The Evidence of the Fixation Points

The writer, employing a tenuous pencil of coloured light, obtained by means of a microstimulation apparatus, for the purpose of exploring the sensitivity of the individual retinal receptors, found that a number of small blind spots, or scotomata, were present near the fixation area of the fovea. These appear to be similar to, if not identical with, those previously described both by Franklin (1894) and by Wilbrand (1896). The cause of these scotomata is not known. Selecting one, which was insensitive to red rays, the writer made use of it for ascertaining the positions in the retina of the fixation points for rays of varying wavelength. Now it is clear that on Young's theory three such points would be expected, since it would be impossible for a receptor to be employed for fixing a ray to which it was itself insensitive. On these grounds there should, therefore, be one fixation point for red rays, another for green rays and a third for blue rays.

In fact, when the matter was investigated by the writer, seven different fixation points were discovered, for the following spectral regions: red (with centre at about 0.66μ),

orange (0.62μ), yellow (0.58μ), green (0.54μ), green (0.52μ), blue green (0.50μ), blue violet (0,46μ) and violet (0.42μ). The positions of these fixation points in the visual field are given in Fig. 3.



Plan of Fixation Area

Fig. 3

Now it is clear that this evidence supports a polychromatic arrangement of the retinal receptors more strongly than it does a trichromatic one, but before accepting this conclusion, one must consider evidence both for and against it. In the latter category come the recent researches on eye movements performed by Lord and Wright (1948) at the Imperial College.

They appear to have shown (*a*) that the eye is constantly performing small irregular rotations, so that the gaze is now directed upwards and now downwards; now to one side and now to the other; (*b*) that in addition it performs rapid flicks, lasting 0.02–0.03 sec., which consist of momentary angular movements of amplitude ranging from about 3 to 14 minutes of arc. These flicks recur once or twice every second; (*c*) That after each flick the eye may take up a slightly new fixation direction, which is displaced some few minutes of arc from the previous direction.

With regard to irregular eye rotations, these are well known in medical ophthalmology, in the condition called nystagmus. A nystagmus may be coarse or fine; may be fast or slow; may be temporary, or constantly present; and so on. A nystagmus is usually incompatible with a high degree of visual acuity (such as 6/4 on Snellen's scale).

Rapid flicks of the eye often occur in all subjects, in some more frequently than in others. When they take place the subject is usually fully aware of them.

A change in fixation direction is quite possible, particularly when the eye is not fully light-adapted or where there is too large an object for precise fixation. But there is another explanation of the facts observed by Lord and Wright, namely, that the eye rotates not about a fixed point like a ball within a metal cup which fits it closely, but about a point which varies in position from time to time. Thus the eye does not form part of a ball-and-socket joint, like the hip joint, but resembles a ball in a sling, the latter in the case of the eye being composed of a thin sheet of fibrous and smooth muscle tissue, called Tenon's Capsule. This is immersed in a mass of semi-liquid fat which fills the posterior part of the bony orbit. Owing to these anatomical arrangements, the eyeball is readily mobile in a sideways direction, so that it can be deflected to one side or another by the slight touch of the end of a finger. Under these conditions the reflections at the corneal surface (the front surface of the eye), such as those recorded by Lord and Wright, may show displacements which are due solely to these sideways displacements of the eyeball and not to changes in the position of the point of fixation on the retina.

In support of the results on the various fixation points for the various colours described earlier in this section are the results of Thomson, who found that differences of wave-

length occur between two neighbouring fields which match in hue, and that these vary with the average wavelength of the two fields. He concluded that the differences in hue which he observed might be due to the position of the fixation point being a continuous function of the wavelength. This conclusion is in agreement with the varying positions of the fixation points for rays of different wavelength, found by the author, which have been described above in this section.

VI. The Subjective Colours

The two properties of the human eye dealt with in sections 4 and 5 above must be ascribed principally to the receptors themselves, because both the retinal direction effects and the many fixation points indicate the existence of many types of receptor which possess specific colour properties. Two alternative possibilities remain to be discussed: (*a*) that all the receptors responding to long wavelength rays may be connected to one visual centre of the brain; those responding to medium wavelength rays may be connected to a second visual brain centre; while those for short wavelength rays when they are stimulated may send nerve impulses to a short wavelength centre. Thus there may be, as it were, polychromatism of the retinal sense organs but trichromatism of the brain, and therefore of the organ of vision as a whole. Alternatively, (*b*) that the polychromatism extends to the brain itself, so that the organ of vision is wholely polychromatic. The evidence to be considered in this, and in the sections which follow, shows clearly that the latter is the more likely of these two rival views.

Let us suppose for purposes of argument that the visual centres in the brain are trichromatic; then if by means of a very tenuous beam of light single receptors could be stimulated, the sensations aroused by such microstimulation would be of three varieties only: a long wavelength variety, a medium wavelength variety and a short wavelength variety. Thus suppose a beam of white light to fall on a yellow

receptor and be connected by a nervous pathway to the long wavelength part of the visual cortex; if the stimulus be adequate, the response of this receptor will lead to the sensation which is normally associated with long wavelength rays. Now since, ex hypothesi, there are only three brain centres, no matter what the spectral content of the incident light stimulus may be, and no matter what the spectral response of the receptor on which it falls may be, only one of these three sensations, and no more, can result from such a stimulus; it must be a sensation which is normally associated with the incidence of long wavelength rays, or one associated with the incidence of medium wavelength rays, or one associated with short wavelength rays. There is no other possibility. If on the other hand there is polychromatism of the higher centres as well as of the retinal receptors then we expect to find more than three specific sensations.

Experiments on these lines were performed by Holmgren (1889) who was the discoverer of the wool test for colour blindness. In one experiment he used monochromatic yellow light (0.5893μ) which was separated from the light of a sodium flame by means of a spectroscope. When a narrow beam of light from this source was caused to move over the surface of the retina, Holmgren found that it might appear either red, green or colourless. When he used monochromatic blue light, in a similar way, he found that it might appear either green, blue, violet or colourless. Thus he discovered four specific colour sensations: red, green, blue and violet. He says in his paper that experiments were also performed in which white light was used as the stimulus 'and that these will be described in the next chapter.' But the writer has so far been unable to trace any description of these further tests.

When the writer employed this method in 1945–6 he was unaware of Holmgren's experiments, and he used in turn not only rays isolated from all parts of the visible spectrum but white light as well. The specific colours which were found were pink, orange and pale blue-green. Violet was

never found when this simple technique was employed, so that the range of colours found by him was less extensive than that found by Holmgren. If, however, a second (conditioning) light source was also presented to the fovea, at the same time as the test light source, then a much greater range of subjective colours could be obtained, and these could be matched without difficulty by means of monochromatic colour fields of known wavelength. The following results were obtained: red (0.64μ), orange (0.61μ), yellow (0.575μ), green (0.54μ), blue green (0.51μ), blue (0.48μ) and blue violet (0.45μ). A description of the apparatus used by the writer and additional details about the results will be found in a previous communication (1947).

The results of some later confirmatory experiments are to be published elsewhere. It should be pointed out here, however, that the colours observed were always of specific hues which were easily identified in wavelength. Thus orange of 0.61μ and yellow of 0.575μ were observed, but never yellow-orange or orange-yellow of 0.59μ or 0.60μ wavelength, and similarly for the other colours. Thus seven specific colours were found in all. Recently when using rays of short wavelength a violet pink or pale crimson response has been obtained, the maximum effect being produced by rays of about 0.42μ wavelength. This specific colour corresponds not to any one wavelength but to some mixture of red rays with blue rays in which the former predominate.

VII. Reduced Foveal Vision

The term 'reduced foveal vision' has been used by the writer to designate the type of colour vision which is found at the fovea when light intensity is low, or the visual angle is small which the test object subtends at the eye of the observer. In the writer's view the types of vision met with under these two conditions are so nearly alike as to justify a single description.

Briefly the facts are as follows: according to Willmer and Wright (1945) all the spectral colours can be matched by

suitable mixtures of red rays and blue rays; and there are two neutral points, one in the yellow at 0.578μ and one in the violet at 0.41μ. According to the writer, using his 'subjective' method, the red end of the spectrum is replaced by a purplish crimson; the yellow by neutral grey; the green by pale blue-green; the blue-green is unchanged; the blue by bluish grey; the blue-violet by neutral grey.

The method of Willmer and Wright shows that there is dichromatic vision comprising long wavelength and short wavelength components. The writer's subjective method shows the same facts, and also that the long wavelength component is a purplish crimson; that the short wavelength component is a blue-green and that these are conplementary in hue to one another, producing white when mixed in suitable proportions.

It is no simple task to find an adequate explanation of these colour changes on Young's three-colour theory. A tentative explanation advanced by the writer on the basis of this theory was that (*a*) the nervous pathway between the blue receptors and the brain centre for the perception of blue light is disconnected at one of the synaptic relays, thus causing the blue-violet end of the spectrum to be replaced by dark grey or black, (*b*) the nervous pathways between the green receptors and the brain centre for the perception of green is cross-connected to the brain centre for the perception of blue as well, thus causing green to be replaced by blue-green, as experiment shows to be the case. Now since the blue receptors are out of action, yellow and white can no longer be distinguished from one another, since both now stimulate the green receptors and the red receptors and thus lead to sensations of blue, green and red together, that is of white light.

But other changes besides those mentioned above occur; thus with reduction of light intensity 'Sloan's notch' (1928) develops in the luminosity curve in the yellow parts of the spectrum, and with reduction of visual angle a corresponding notch, 'Wright's notch' (1942), develops in a similar

way. Another change is found to occur in the shape of the red response curve in the short wavelength part of the spectrum. In order to account for all these changes four *ad hoc* hypotheses have to be advanced, namely:

1. The cutting of the 'blue' pathway.
2. The linking of the 'green' pathway to the blue centre.
3. The enhancement of the red response at the long wavelength end of the spectrum.
4. The enhancement of the red response at the short wavelength end of the spectrum.

These four additional hypotheses not only give one another no mutual support, but are not supported by any other evidence; thus the three-colour theory finds itself faced with a very serious situation. The polychromatic theory, on the other hand, finds itself confronted with no such difficulties, as will be explained in the next section.

(*to be concluded in the next issue*)

Measuring the Universe

WILLIAM H. MARSHALL

IT is the misfortune of the astronomer that the methods by which he measures the universe are overshadowed by the sensational results he achieves. It is probably true to say that for every layman who knows how astronomical distances are measured, there are at least ten, perhaps even a hundred or more, who know something of the distances themselves. And yet the story of how the results are obtained is infinitely more interesting and instructive than a mere recital of millions piled upon millions until understanding is lost in amazement.

In the strictest sense, there is only one method of *measuring* distances, and that is by laying a graduated scale between the two points whose distance apart is required. All other operations leading to results which we call distances might more properly be called methods of *estimating* distances; and the results of such operations are valid only so long as they have a firm and demonstrable connection with the method of the graduated stick. Nevertheless, it is customary and convenient to speak of measuring the universe, and the methods are such that no confusion is likely to arise.

The limitations of the direct method of measuring lengths seldom concern us in everyday life, because we rarely have occasion to measure distances which are either very big or very small, and so the ordinary scale is quite accurate enough. When, however, we begin to measure lengths much bigger or smaller than ourselves, we soon find ourselves in difficulties. The methods of overcoming these difficulties are quite different for the very big and the very small; but it is worth while to mention that some methods of measuring great

distances in space depend upon the measurement of small ones in the laboratory. The technique of measuring these small distances will have to be taken for granted here.

The first extension of measurement to long distances is the surveyor's method of triangulation. There are obstacles which prevent the direct measurement of the distances AC and BC, for example. In the first place, these distances may

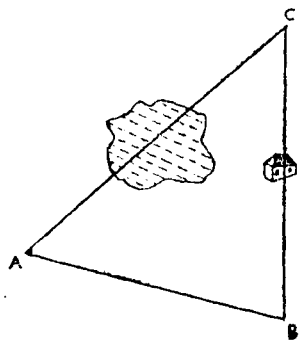


Fig. 4.

be too long for accurate direct measurement; and secondly, there may be physical obstructions such as hills, lakes, towns, and so on. To meet these difficulties the surveyor makes the most accurate possible measurement of the line AB, and then measures the angles at A and B, so that he then knows three elements of the triangle ABC. By the use of elementary trigonometry, he can then calculate the lengths of the sides AC and BC. Since angles can be measured with a high degree of accuracy, and since there are no experimental errors in the trigonometrical formulae, this method gives the required distances with an accuracy limited only by that of the measurement of the base-line AB.

By use of this method, it is possible to cover a whole country with a network of triangles which gives the distances between particular points to an accuracy of as high as an inch in twenty miles. By combining this process with astro-

nomical observations, it is possible to measure the diameter of the earth.

At two points a considerable distance apart on a meridian of the earth, the astronomer measures the angle between the zenith and the pole, as shown, and the difference between

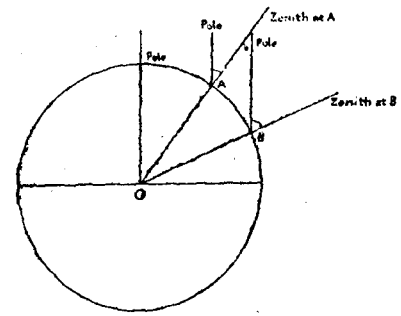


Fig. 5.

these two angles is the difference of latitude of the two stations ($\widehat{AOB}$). By triangulation, the distance between A and B is found, and hence the circumference, which is the distance corresponding to a latitude difference of 360°. If the earth is assumed to be a true sphere, the diameter can immediately be worked out from this. A refinement of the method, in which the distances corresponding to an angle of one degree at various parts of the earth is measured, shows that the earth is not a perfect sphere. It is, in fact, somewhat flattened, measuring 7,900 miles from pole to pole, and 7,927 miles across the equator.

From Fig. 4 it is clear that the farther away C is from A and B, the smaller will be the angle subtended at C by the base-line AB. If C is taken successively farther and farther away, this angle will finally become too small to be accurately measurable with the scales of a theodolite. If we wish to measure AC, it is then necessary to extend the base-line beyond B until it subtends an accurately measurable angle at C. This problem does not usually arise in surveying, because we can fill in an inconvenient space by means of sub-

sidiary triangles. But when we wish to measure off the earth, we have no choice but to extend our base-line in any way we can.

The Moon

For the determination of the distance to the moon, the base-line taken is the distance between two widely separated observatories, which for simplicity we shall assume to have the same longitude. The angle between the moon and the zenith is measured at the same time at the two observatories.

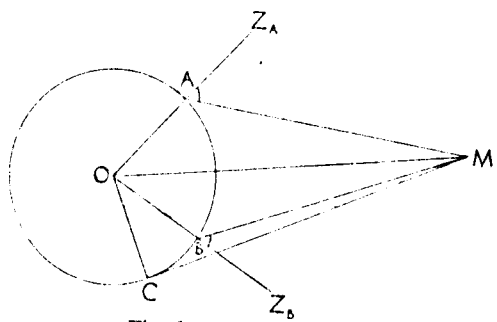


Fig. 6.

We then know the following elements of the quadrilateral OAMB: $\widehat{OAM}$ and $\widehat{OBM}$ from the observations; $\widehat{AOB}$, the known latitude-difference of A and B; OA and OB, which are radii of the earth. These data enable us to solve the quadrilateral, and in particular to find the distance OM. In practice, the procedure as described is modified, so that the observatories need not be on the same meridian, nor need the observations be simultaneous. This makes the calculation of the results more complicated, but the method is still fundamentally the one described.

It is obvious from what has been said that the relative distances of a series of objects could be specified by angles instead of a linear measure. The angle in each case is the angle subtended at the object by a particular base-line, and, so long as the same base-line is used for all the objects, this

Plate 3. *Daphnia* with haemoglobin in blood and eggs.

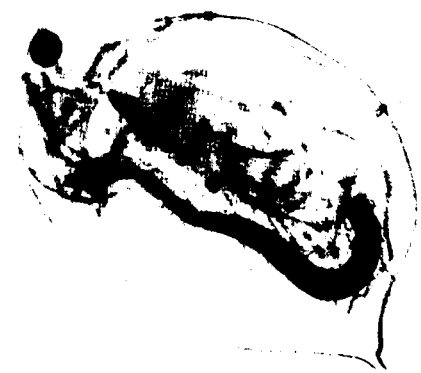


Plate 2. Pale *Daphnia*, with only traces of haemoglobin in its blood.

(See the article on 'The Blood of Water Fleas' by Professor H. Munro Fox, F.R.S.)

Plate 4. Apus, a primitive crustacean with haemoglobin in its blood. (p. 21)

angle, called the *parallax*, is a satisfactory index of distance. Since in most astronomical work it is the angle which is measured, it is generally more convenient to speak of the parallax of an object, rather than its distance. In the case of the moon, the base-line chosen is the equatorial radius of the earth, drawn to the point where the moon is on the horizon (OC in Fig. 6). The angle in question is then $\widehat{OMC}$. This is called the moon's equatorial horizontal parallax, and its mean value is 57′ 2″.7, corresponding to a mean distance of 238,857 miles. The distance varies, as the moon moves around the earth, from 221,463 miles to 252,710 miles.

The Solar System

The possibility of measuring the moon's distance in this way depends on two facts: first, the distance is small enough for the angle AMB to be accurately measurable, and second, the surface of the moon offers many features on which instruments may be sighted with certainty. This is not the case, however, with the other major members of the solar system. To obtain accurate parallaxes for them it is necessary to have recourse to a modification of this method, which depends on the observed motions of the planets.

If we wished to determine the distance of a *fixed* object P, we could let the earth's motion in its orbit from E_1 to E_2 provide a base-line, by measuring the position of P at times separated by a measured period. In the case of the planets,

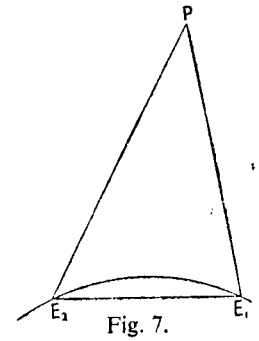


Fig. 7.

this simple scheme is upset by two factors. First, we do not know the distance $E_1 E_2$ until we know the earth's distance from the sun; and second, the planets are not fixed.

Although we do not know the radius of the earth's orbit at this stage, we do know its shape from simple observations. The apparent size of the sun on any given day, as seen from the earth, clearly depends on its distance from the earth on

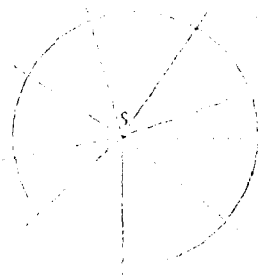


Fig. 8.

that day. Taking a point S to represent the sun, we draw from it a set of lines, so that the angle between any two represents the time between the corresponding observations (360° representing one year). We then lay off on these lines distances proportional to the observed apparent sizes of the sun, and the figure formed by drawing a smooth curve through the points so obtained will be a plan of the earth's path around the sun. In this way, we find that the earth's orbit is an ellipse of small eccentricity. Knowing that the whole circumference of the ellipse is traversed in 365¼ days, we can then calculate the arc traversed in any given time, and also the corresponding chord, i.e., in Fig. 7, the distance E_1E_2, *in terms of the earth's mean distance from the sun.* If we did this for all the planets, we should be able to find all their distances as multiples or fractions of the earth's distance from the sun, and so we should be able to draw a map of the solar system, lacking only the scale of distances. Further-

more, the determination of any one distance on this map would enable us to find all the others.

The planets, however, are not fixed, and so the method just described could not be used. In Fig. 9 we see that, while

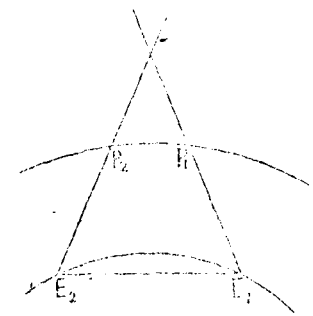


Fig. 9.

the earth moves from E_1 to E_2, the planet itself moves from P_1 to P_2, and so the point of intersection of the two lines of sight is not at the planet but at some other point in space. All we require, however, is that the planet should be at the same position at the times of the two observations; it is not necessary that it should stay there during the intervening time. If, therefore, we know how long the planet takes to go round the sun (its Sidereal Period), we can space our observations at that interval, and the difference between that period and one year (the earth's sidereal period) is the time between the earth's positions E_1 and E_2. This will clearly give us the information we require, and all we need now is a method of finding the sidereal period of the planet. Since this is the time required for the planet to return to a given place among the stars, *as seen from the sun*, it is clearly not possible for us to observe it directly. But we can find it as follows.

Although the sidereal period of a planet cannot be observed, we can observe its Synodic Period, which is the time it requires to return to the same position relative to the sun *as seen from the earth.* Suppose this time is S days, and

that the sidereal periods of the planet and the earth are P and E days respectively. Then the earth traverses an angle of 360° in E days, the planet in P days; and the earth gains or loses 360° with respect to the planet in S days. Therefore, taking daily motions, we see that the earth moves 360°/E in one day, the planet 360°/P; and the earth gains or loses 360°/S in one day. But the angle gained or lost by the earth in one day is the difference in the daily motions of earth and planet. Hence $1/S = 1/E - 1/P$ or $1/P - 1/E$, according as earth or planet moves the faster. Since both S and E can be observed, we can easily calculate P. Thus we can space our observations of the planets so that they are observed always at the required parts of their orbits.

As explained above, this enables us to make a map of the solar system, all of whose distances are known in terms of the distance of the earth from the sun, a distance which is called the Astronomical Unit. The value of this unit, and so of all the other distances on the map, is found by the determination of the parallax of some member of the system.

The most accurate value will be obtained from the determination of the smallest possible distance. Nowadays, the body chosen is the minor planet Eros, which may at times approach to little more than thirteen million miles from the earth. This minor planet has the advantages that its distance is small enough to be measured accurately by the method described for the moon, and that the planet presents a disc so small that it is easy to set instruments accurately on it. It was from observations made by observatories in fourteen countries at the close approach of 1930–31 that Sir H. Spencer Jones, the Astronomer Royal, derived the now accepted solar parallax of 8″.79, corresponding to a mean distance of 93,004,000 miles. During the year the distance varies by about three million miles, being least in January and greatest in July.

The application of this scale to our map gives the following distances for the planets from the sun. The distances are given first in astronomical units, then in millions of miles.

Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune	Pluto
.387	.723	1.000	1.524	5.203	9.539	19.191	30.071	39.518
35.09	67.24	93.00	141.7	483.9	887.1	1785.8	2796.6	3675

There are other methods of determining the solar parallax, the best of them being those that depend on gravitation theory and the velocity of light. Since, however, they serve mainly to corroborate the results of the trigonometric method, we need not describe them here.

The Near Stars

When we have found the distance of the earth from the sun, and so the size of the solar system, we have to take another step out into space – this time to the stars. And once more we find our base-line in need of extension. We have just seen that the earth is moving round the sun in a nearly circular path of about 186 million miles diameter. It would be reasonable, therefore, to expect at least the nearest of the stars to show some change of apparent position as we move from one end of a diameter to the other; and so they do. But the change is so small that it was only in 1838 that Bessel first succeeded in measuring it for a star, 61 Cygni.

To use our new base-line, we obviously have to measure the position of the parallax star (i.e., the star whose parallax is being measured) relative to neighbouring stars at intervals of six months. The earth being then at opposite ends of a diameter of its orbit, we should observe an apparent change in the star's position. The angle which is actually measured is that subtended by a diameter of the earth's orbit; and half

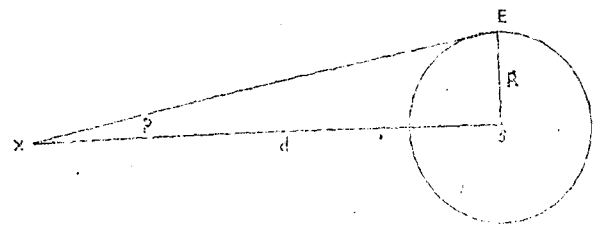


Fig. 10.

of this, i.e., the angle subtended by the radius of the orbit, is called the 'heliocentric' or 'annual' parallax of the star.

In Fig. 10, if S, E, and X are the sun, the earth, and the star, the star's parallax is SXE. Now suppose SX = d, SE = R, and SXE = p. Then d = R/tan p*. Since p is a small angle, we may replace its tangent by its radian measure, so that we have d = 206,265R/p, where p is measured in seconds of arc. Since the earth's orbit is slightly elliptical, R is the mean radius. (206,265 is the number of seconds in one radian).

The motion of the star in the sky, however, is rarely if ever purely parallactic; it is compounded rather of an elliptical motion which is the reflection of the earth's orbital motion seen obliquely from the star, and a straight line motion, which is the proper motion of the star, i.e., its motion relative to the sun. Hence we need not two, but at least three, and preferably many more, observations. These enable us to

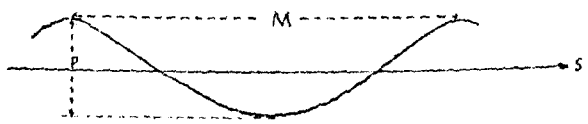


Fig. 11.

draw a picture (Fig. 11) of the star's apparent path. Here, the star is really moving along the straight line S, but it suffers also the apparent displacements caused by the movement of the earth. From the observations we get P, twice the parallax of the star, and M, its proper motion.

But what of the comparison stars? Suppose they, or some of them, are themselves moving? Then most of the work is lost. Fortunately, however, most of the stars in the sky are too distant to show any appreciable motion due to parallax; at the worst, the mean displacement of the comparison stars has been found to be a few thousandths of a second of arc, and in modern work even this can be allowed for fairly accurately.

It was mentioned above that stellar parallaxes are very

*See any elementary textbook of trigonometry for the meaning of tangents and radians.

small. The largest of them, in fact, is only three-quarters of a second of arc, which means that the angle to be measured is in most cases smaller than 1″.

The modern method of finding trigonometrical parallaxes is by photography. A part of the sky containing the parallax star is photographed with a telescope of long focal length. The telescope must be carefully guided, so that the smallest possible image is obtained; the nearer the image is to a point, the more accurately can its position be measured. And if necessary the light of the star is reduced to give an image of about the same brightness as those of the comparison stars. The position of the parallax star relative to half a dozen or so others is measured on up to twenty plates taken at various times, and the parallax and proper motion are calculated from these measures. In this way it is possible to find the parallax of a star down to about a two-hundredth of a second of arc, though the smallest of the parallaxes are not very accurate.

The applicability of the method is again limited by the ultimate impossibility of measuring accurately an angle less than a few thousandths of a second of arc. But it enables us to survey that part of space where parallaxes range from 0″.75 down to 0″.005. In terms of distance, this means from about 25 million million miles to about 3,000 million million miles from the sun.

Units

At this point the question of units becomes serious. It will be remembered that even in dealing with the solar system, astronomers found it convenient to use a new unit of distance, the Astronomical Unit, equal to 93 million miles. One reason for the adoption of this unit is that it is no more convenient to express the distance from, say, the Earth to Neptune, in miles than it is to express the distance from London to Glasgow in inches. Now, the lower and upper limits mentioned for the trigonometrical method are about 250,000 and 30 million astronomical units respectively, so here again we need new units of distance.

The unit most commonly used in non-technical literature is the light-year. This is the distance travelled by light in one year at a speed of 186,000 miles per second, and it is equal to nearly six million million miles. Using this unit, the distance of the nearest star is about 4¼ light-years, and the limit of the direct method is about 650 light-years. This unit has the incidental advantage of reminding us of the temporal as well as the spatial aspect of the universe. Light takes *x* years to reach us from a body x light-years away, so that we are really seeing the past condition of the source. In similar units the moon is rather less than 1½ light-seconds from the earth, and the sun some eight light-minutes. In more technical work, the unit commonly used is the parsec, which is the distance at which a body would have a parallax of one second of arc. One parsec is equal to 3·26 light-years, or about 19 million million miles. For longer distances still there are the derived units, the kiloparsec and the megaparsec, a thousand and a million parsecs respectively. In parsecs the limits of the direct method are about 1·3 and 200 parsecs.

By the trigonometrical method it is possible to measure the distances of a few thousand stars; but it will be remembered that the method depends on the use of stars too far off to be themselves measured in this way. And in fact the few thousand stars whose distances are measured directly are a very small fraction of the many millions in our own galaxy. The number is slightly increased by the dynamical parallaxes derived from the study of double stars, and by the statistical studies of the movements of various types of stars. But for any significant extension of our 'measuring rod' we have to turn to quite other considerations, and to methods which have very little appearance of dealing with distances at all.

Distant Stars

The physical, as distinct from the trigonometrical, methods of finding stellar parallaxes depend ultimately on a simple idea. The apparent brightness of a source of light varies directly as its real brightness, and inversely as the square of its distance from the observer. If, therefore, we know the distance and the apparent brightness of a source of light, we can calculate how bright it really is; and, conversely, if we know both its intrinsic and its apparent brightness, we can find its distance. Now, one of the most readily observed characteristics of the stars is the difference in brightness between one and another. These differences have been observed over many centuries, and the stars have long been catalogued in classes according to their apparent brightness. For the naked-eye stars, the brightest are said to be of the first magnitude, the faintest are of the sixth magnitude, and the intermediate ones are grouped between so that the ratio of brightness from one class to the next is 2.512. The magnitude scales now in use include zero and negative magnitudes for the brightest objects, and the faintest stars photographed with the 100-inch telescope at Mount Wilson are of about magnitude 21.

The brightnesses thus catalogued are, however, only apparent. Thus, two stars may appear equally bright, either because they have the same actual luminosity or, as is more likely, because the brighter is farther away from us than the fainter. So, in order to compare the light-outputs of the stars, we introduce the idea of Absolute Magnitude. The absolute magnitude (M) of a star is defined as the magnitude it would have *if it were situated at a distance of ten parsecs from the sun.* It is connected with the apparent magnitude (m) by the formula $M=m+5+5.\log p$, where p is the parallax. Thus, whenever we have determined the parallax of a star, we can calculate its absolute brightness from its apparent brightness; and conversely, if we can in any way find the absolute magnitude of a star, we can find its parallax. Before proceeding farther, it may be interesting to note the apparent and absolute magnitudes of some well-known stars.

	Sun	Sirius	Capella	Antares	Vega	Betelgeuse	Castor
m	—26.7	—1.58	0.21	1.22	0.14	0.92	1.58
M	4.85	1.3	—0.6	—4.0	0.6	—2.9	2.2

The most striking feature of this list is the comparatively humble figure for the absolute magnitude of the sun, which at the standard distance of ten parsecs would be just comfortably visible to the naked eye on a clear, dark night. This emphasises the fact that the sun's outstanding brilliance is due entirely to its nearness.

How, then, do we find the absolute magnitude of a star whose distance is not already known? The first method is by examination of its spectrum. When light is passed through a prism or grating, it is spread out into a spectrum, a band of light showing the rainbow colours red, orange, yellow, green, blue, indigo and violet, from one end to the other. This spectrum, in the case of a star, is crossed by dark lines which convey to the astronomer information about the temperature, composition, and velocity of the star. The spectrum of any star can be classified, according to the lines in it, in one of the groups O, B, A, F, G, K, M, R, N, S, and within the group in one of ten sub-groups. Some typical spectra are shown in Plate 17.

Stars of the same spectral type are not, however, all of the same size and luminosity; and in 1914 Adams and Kohlschütter discovered that there were systematic differences between the spectra of bright stars (giants) and faint ones (dwarfs) of the same spectral class. Plate 18 shows two spectra illustrating the differences. The relative intensities of some lines are quite different in the two spectra; and in particular it is found from the study of nearby giants and dwarfs whose parallaxes are known that the ratio of the intensity of a line due to ionised strontium to that of a line of iron provides a sensitive means of estimating absolute magnitude.

Thus, if we can photograph the spectrum of a star on a scale big enough to enable us to measure the intensities of these two lines, we can find its absolute magnitude, and hence its parallax and distance. The method is limited to comparatively nearby stars by the need for large-scale spectra. But it reaches farther than the direct method, and

the number of parallaxes obtained by it is far in excess of the number obtained trigonometrically. Also, for the smaller parallaxes, the spectroscopic method is more accurate. The probable error of a single determination is about fifteen per cent, and it is the same for small parallaxes as for large ones. In the trigonometrical method, on the other hand, the probable error is about 0″ .005 and is the same size for all parallaxes, so that the percentage error increases as the parallax decreases. Thus, for parallaxes less than about 0″.05, the spectroscopic method is superior to the trigonometric.

The Universe

So far, the methods described have been applicable only to stars. The next method is one which enables us not only to extend our measurements to stars far more distant than any otherwise attainable, but also to measure the distances of external stellar systems comparable in all respects to our own galaxy. This method uses the characteristics peculiar to the type of stars known as Cepheid variables. These stars, so called because the first one studied was Delta Cephei, are of variable brightness, the magnitude of each varying in a quite definite way, and with a definite period of change.

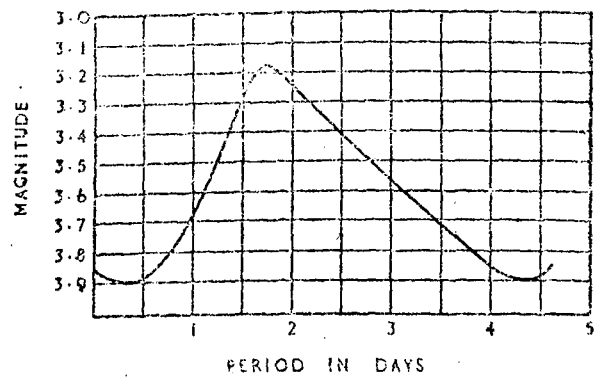


Fig. 12.

Fig. 12 shows the kind of curve obtained when the brightness of a Cepheid is plotted against the time. The fairly rapid increase in brightness followed by a comparatively slow decline, the whole cycle being repeated indefinitely, is typical of this kind of star.

The peculiar feature of these stars which enables us to use them as indicators of distance, was discovered by Miss Leavitt of Harvard in the early years of this century. While studying the Cepheids in the Magellanic Clouds (Pl. 20), Miss Leavitt found that in each Cloud separately the mean apparent magnitudes of the Cepheids were closely related to their periods of light-variation. Thus if either the mean brightness or the period were known, the other could be inferred. Now, if we consider either of the Clouds alone, we are dealing with a collection of stars at appreciably the same distance from the sun, since the size of the Cloud is very much less that its distance. So what is true of the apparent magnitudes of these Cepheids must also be true of their absolute magnitudes. And so a relation was found which made it possible to derive the absolute magnitude of a star of this type from the period of its variation ... provided the zero-point of the scale could be found. This was done by investigating Cepheid variables in our own galaxy. Fortunately there are some whose parallaxes can be found by other methods, and whose absolute magnitudes can therefore be found and used to calibrate the new relation.

The Cepheid method is used as follows. When a star has been found to be variable, its magnitude is measured many times at frequent intervals, so that its light-curve can be drawn. If the variation proves to be of the Cepheid type, we then get from the curve the mean apparent magnitude and the period. From the period-luminosity relation we then find the absolute magnitude, and so, knowing both m and M we can find the distance. This method is available up to any distance at which we can see a Cepheid; and fortunately they are very bright stars. They enable us to measure distances up to about a million light-years, and that with a

percentage error less than in the measurement of a thousand light-years by the trigonometric method.

The number of Cepheids is not very large, but they have enabled astronomers to extend their measurements far beyond the limits of our own stellar system. In 1924 Hubble at Mount Wilson discovered several Cepheids in the Andromeda nebula (Pl. 19). When these were studied, they proved to be at distances of the order of 700,000 light-years. The nebula was thus proved to be an independent system similar to our own galaxy. When Cepheids were subsequently discovered in other spiral nebulae, the way began to open for the extension of distance-measurements to their present limits.

The way was further cleared by the study of the nebulae near enough for individual stars to be recognised in them. In addition to Cepheids, it is possible to recognise two important types: novae and blue giants. Novae are stars which suddenly flare up with explosive violence from comparative faintness to very great brilliance. Some twenty or thirty are observed every year in our own and other galaxies and the study of the nearest of them shows that they reach on the average an absolute magnitude of about —5; they can therefore be used as rough indicators of distance. The brightest giants are found to be about 50,000 times as bright as the sun, and so they also can furnish a rough but useful measure of distance.

The distances derived by means of Cepheids, novae, and blue giants serve to calibrate a new index of distance in the spectra of the nebulae. The lines in any spectrum have quite definite positions, and they are displaced from these positions, in terrestrial experiments at least, only by movement of the source of light relative to the observer. While the spectrum of a star is being photographed, that of a light in the observatory is also photographed on the same plate, so that the positions of the lines in the star's spectrum may be checked. If they are found to be displaced towards the red end of the spectrum, we believe that the star is moving

away from us; if the displacement is towards the blue the star is moving towards us. In each case, the amount of the displacement is taken to indicate the relative velocity.

Now, in the case of the nebular spectra, it is found that, with a few exceptions among our nearest neighbours, the lines are displaced towards the red; and the amount of the displacement is greater the farther the nebula is from us. From the spectra of nebulae whose distances have already been found, a relation is found between the red-shift and the distance. Since, as already mentioned, a displacement of the spectral lines normally indicates movement, the relation is called the velocity-distance relation, and the displacements are usually referred to in terms of the corresponding speeds. Fig. 13 shows some spectra illustrating the red-shift, and the corresponding direct photographs of the nebulae show the increasing faintness which suggests increasing distance. From spectra obtained with the 100-inch telescope we can estimate distances up to about 250 million light-years.

This is a very gratifying as well as a very surprising consequence of our spreading a faint white light out into an even fainter band of colour. It raises, incidentally, one rather interesting question: we may not be reluctant to believe that some galaxies are so far away that the light reaching us now left them 250 million years ago, but are we so ready to believe that they are travelling away from us at some 25,000 miles per second? May there not be some other cause reddening the light in its journey through space? Astronomers are not agreed on the answer to this question, which, in any case, does not really concern us here. But on one thing they are agreed: no matter what the cause of the red-shifts, these displacements of the lines in nebular spectra provide a reliable scale for the measurement of distances at which no individual stars can be detected.

Beyond the limit of the velocity-distance relation, the measurement of distances becomes largely a matter of statistics. It is fair to assume that the average luminosity of

Spectra of Extra-Galactic Nebulae.

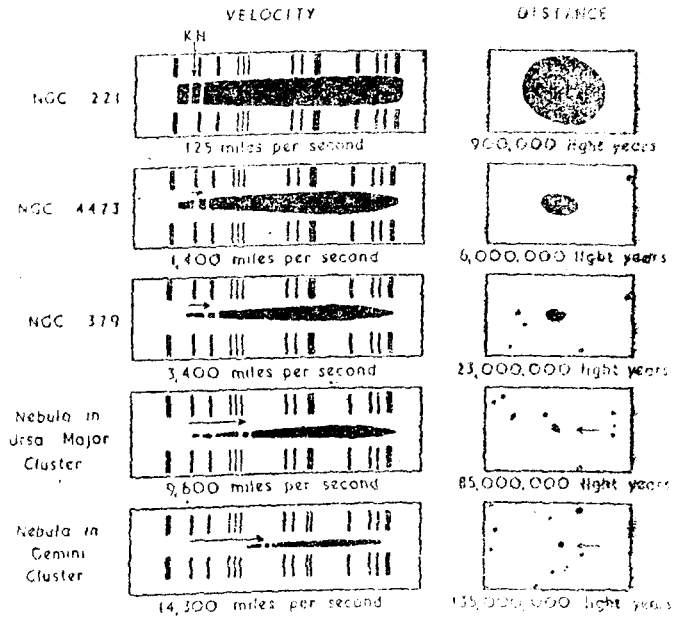


Fig. 13.

The arrows above the nebular spectra point to the H and K lines of calcium and show the amounts these lines are displaced toward the red end of the spectra. The comparison spectra are of helium.

The direct photographs (on the same scale and with approximately the same exposure times) illustrate the decrease in size and brightness with increasing velocity or red-shift.

NGC 4473 is a member of the Virgo cluster and NGC 379 is a member of a group of nebulae in Pisces.

(Redrawn from Hubble's *The Realm of the Nebulae*, by courtesy of the Yale University Press).

the nebulae beyond that limit will be about the same as for those within it. An estimate of distance follows from apparent brightness: an estimate which is likely to be very inaccurate for individual objects, but which, statistically, will give an accurate picture of the distribution in space of

the millions of galaxies out to the present limit of about 500 million light-years.

What of the future? The 200-inch telescope will extend the boundary still farther – to about a thousand million light-years, where we may expect some interesting things to be found. But it will also extend the scope, and the accuracy, of the velocity-distance relation; and it will make possible closer and more extended study of individual stars in individual galaxies. Thus, while it will place new stories on the very top of this edifice, it will at the same time strengthen the parts of the structure which support them.

Nearer home, new telescopes and improvements in photographic emulsions will facilitate deeper study of single stars in our own galaxy; and deeper theoretical insight into the nature of stars may well provide new means of estimating absolute magnitude. But the whole structure will continue to depend ultimately on the 'old-fashioned' astronomy which carries the surveyor's chain of triangles out into space, and so links the galaxies with the measuring rod.

Aspects of Crystal Chemistry

PETER R. ROWLAND

What is a crystal?

Most people imagine a crystal as an object with a highly defined geometrical shape such as a cube or an octahedron or a hexagonal prism, etc. The external shape is not a fundamental property of a crystal, however, but a secondary manifestation of the fact that the atoms or molecules of which it is composed are arranged in a regular geometrical pattern in space. Figure 14 shows a layer, A, of spheres, representing atoms in closest contact, with another layer, B

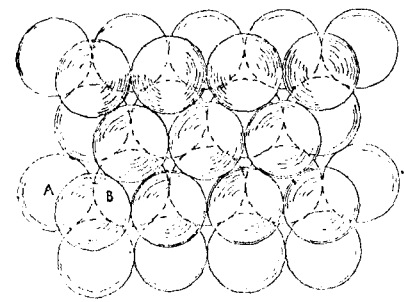


Fig. 14.
Layers of spheres in closest contact. Lower layer, A (unshaded). Next layer, B (heavily shaded).

(shaded) on top, each of the top ones fitting in the dimple between three of the lower ones. A third layer, C, can go on in either of two ways. Its atoms can be over either the atoms or the dimples of the first layer A. In the case of copper crystals, for instance, they are over the dimples and the layers repeat ABC, ABC, etc., throughout the crystal. A

model can therefore be made by adding layers according to this scheme until we have formed a large pile.

Now suppose we take a knife and cut through this pile of spheres with it parallel to any given plane. Remember that we cannot cut through atoms but must simply split them apart from one another. It will be found that the atoms in the surfaces exposed are arranged in a geometrical pattern and that the pattern is always the same, no matter where the cut is made, provided it is made parallel to the same plane all the time. For instance, if our cut is parallel to the layers ABC, etc., it will expose one of those layers and the atoms in the surface will be arranged in rows parallel to the sides of an equilateral triangle. For cuts in certain other directions the atoms will be arranged like the top layer of eggs in a crate, i.e. in rows, intersecting at right angles. In whatever direction we make our cut, the atoms in the exposed surface will be arranged in a definite pattern depending on the directional relation of the cut to the layers ABC, etc.

Consider building up a crystal model from table tennis balls stuck together in layers as described, till they form a mass about 10 feet high. Now suppose we shape this mass into a sphere about 10 feet in diameter by gently removing balls from the outside. The surface of the sphere may be regarded as made up of an infinite number of small planes which will cut the mass twice, parallel to every possible direction, i.e. one plane at each end of every possible sphere diameter. In other words, over any small area the surface of the sphere will be to all intents and purposes a plane with the atoms in it regularly arranged.

Suppose we now examine the surface and compare the patterns found in various parts. We shall find six places with the egg crate pattern and they will be arranged at the ends of three diameters which are mutually perpendicular, i.e. at the points where the sphere would touch if it were put into a huge cubic box that just fitted, with the diameters concerned parallel to the edges of the box. In other words, the arrangement is found to have 'cubic symmetry'.

This is explained in Fig. 15a. The regions such as X, B, C, etc. on the sphere expose the egg crate pattern, and the planes which contain this pattern are parallel to the sides JKLM, LMNO and OPKL of the cube.

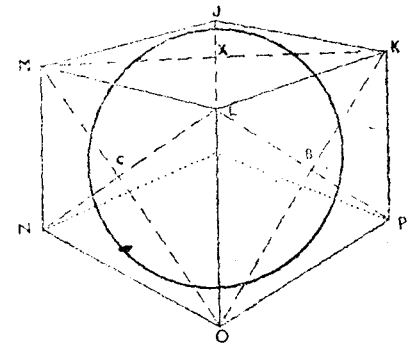


Fig. 15a.

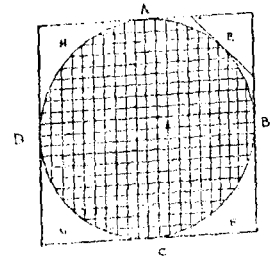


Fig. 15b.

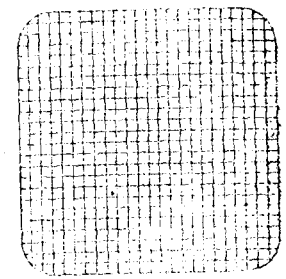


Fig. 15c.

Figure 15b represents a section through the sphere and box in the plane of B and C. The lines represent the layers of atoms with the egg-crate formation, parallel to two planes at right angles, i.e. to two sides of the box. There is another set intersecting these, and parallel to the plane of the paper, i.e. the top of the box. Such an arrangement is called a lattice. At the points of contact with the box such as ABCD it is seen that these planes actually form part of the surface.

Of course real atoms are so small that in any crystal large enough to handle there would be millions of layers.

To sum up, a crystal is a regular arrangement in space of repeating units. If we cut it in any plane we shall expose a surface containing atoms arranged in a geometrical pattern. Such surfaces are called crystal planes. In a simple symmetrical crystal any given pattern may occur in several sets of planes, each parallel to one of the faces of a simple polyhedron, e.g. in the 'cubic close packed' arrangement of spheres just described, the egg crate pattern occurs in three sets of planes parallel to the faces of a cube (cubic planes), the triangular pattern occurs parallel to the four planes which would cut off the corners of the cube (octahedral or 'cube corner' planes), there are six identical sets of planes parallel to the cuts which would bevel off the edges of the cube at 45° (dodecahedral or 'cube edge' planes), and so on.

Where a particular plane forms part of the surface of a crystal it is called a crystal face. For instance, in Figure 15b the spherically-shaped crystal exposes cubic faces at A, B, C, and D, dodecahedral faces at E, F, G and H, and so on. Plate 8 shows this arrangement of close packing of atoms, cut so as to expose cubic faces. It is in fact in the shape of a cube, though it could equally well have been in the shape of a rectangular brick. All that matters is that the surfaces exposed shall be parallel to the appropriate planes. Plate 10 shows a further cut made to expose an octahedral face and Plate 9 a cut made to expose a dodecahedral face. Note the egg-crate pattern with atoms in diagonal rows on the cubic faces, the triangular pattern on the octahedral face and the very open pattern of horizonal rows in the dodecahedral face.

The direction perpendicular to a cubic plane is called a cubic direction, that at right angles to an octahedral plane is called an octahedral direction, etc.

The crystal described above is a particularly simple one. Where the units are not simple spheres all of one size, the

possible number of patterns becomes enormous. However, the principles involved remain the same.

We are now in a position to consider why the crystals of salt, sugar, native copper, etc., display such definite shapes. Suppose we start with our spherical crystal and grow it into a larger crystal by depositing more atoms on it. In practical cases it is found that the rate of deposition is not the same at all faces, but that some grow faster than others. If the dodecahedral ('cube edge') faces EFGH grow faster than the cubic faces the shape will become like that in Figure 15c, i.e. the crystal tends to acquire a cubic shape. Rapidly growing faces tend to obliterate themselves and the slowest remain to give the crystal its shape. The particular shape exhibited by a substance growing under given circumstances is called its crystal habit. When crystals are growing from solution, it is found that the presence of traces of other substances will profoundly influence the relative rates of growth of the various faces, so that a substance may have its crystal habit modified by an impurity. The final shape is usually just a result of the mode of growth. A crystal can in fact have any external shape. What determines that it is a crystal, is the fact that all the atomic or molecular units of which it is composed are arranged in a regular and continuous three-dimensional pattern.

Methods of Growing Crystals

There are four principal ways of growing large single crystals:

(*a*) By depositing fresh material on to a small seed crystal from the vapour.
(*b*) By doing the same from solution.
(*c*) By allowing the molten material to solidify slowly.
(*d*) By annealing solids at high temperatures.

The first two methods are not very convenient since one has little control over the shape of the final crystal and it is difficult to prepare large crystals free from flaws. Method (c) is the most useful. A typical example of its use is shown

diagramatically in Figure 16 for the growth of metal crystals. The crucible has a conical bottom to its cavity and, when the furnace is raised slowly, the tip of the cone cools first and the first crystal forms here. As the furnace continues to rise, this crystal grows upwards until the whole melt has been converted into a single crystal ingot with an external shape fitting the crucible cavity.

Furnace Moves Slowly Upwards

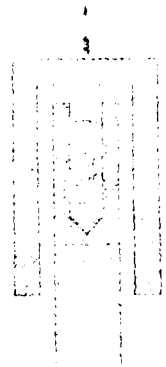


Fig. 16.

An alternative is to fix a suitable 'seed' crystal in the surface of the melt with its crystal planes orientated the way you want them in the final crystal (e.g. you may want one set of cubic planes perpendicular to the length of the ingot and the other two parallel to it). The furnace is then lowered gradually and the crystal grows from the top down. Of course there are plenty of complications. Metals must be protected from oxidation by enclosing the whole apparatus in a vacuum or an atmosphere of inert gas. Furnace temperature control must be good and the rate of cooling adjusted very carefully by making the furnace travel very slowly (about 1 cm. per hour or less). By methods such as this, it is possible to prepare synthetic crystals of rock salt and similar minerals of the highest optical quality and about as big as a man's head. This is now a regular industry.

The idea behind the fourth method is to grow one of the minute crystallites of which a piece of metal (or other solid) is made up at the expense of its neighbours until it 'occupies' the whole piece. This can be done by straining the metal (e.g. stretching it one per cent in length), then keeping it at a temperature near its melting point for several hours. The strain will warp the lattice in most of the crystals and this makes them unstable. At the high temperature the atoms vibrate about their mean positions with large amplitudes. Where a relatively unstrained lattice shares a boundary with a strained one, atoms will jump right out of the latter into the former more stable lattice and so this crystal grows at the expense of its neighbours. Very often one crystal will engulf all the rest, giving a single crystal specimen. Professor E. N. da C. Andrade has developed an ingenious method for growing single crystal iron wires. The wire is heated by passing a current through it. It is put into a subsidiary furnace which maintains a temperature gradient from one end to the other and allowed to cool by decreasing the current. Iron changes its crystal structure at 900°C so that a single crystal of the low temperature form grows along the wire from the cool end.

Unfortunately, annealing methods do not work for all metals.

What holds a Crystal together?

There are two principal kinds of bond between atoms. In one, an electron hops out of one of the two linked atoms into the other, leaving the first with an electron deficiency (positive charge) and giving the second a negative charge. The atoms are then attracted to each other because they carry opposite charges. It may be asked why the positive charge on the first atom (or rather ion, as a charged atom or molecule is called) does not pull the electron back into itself. The answer is that when we work it out we find that it would require more energy to do this than is required for the separation of charge to remain. For example, when the

metal sodium reacts with the gas chlorine this energy is given out as heat and radiation and so lost from the system. When the product is cooled it is found to be crystalline and made up of positive sodium and negative chlorine ions packed together in the three-dimensional chessboard pattern shown in Figure 17.

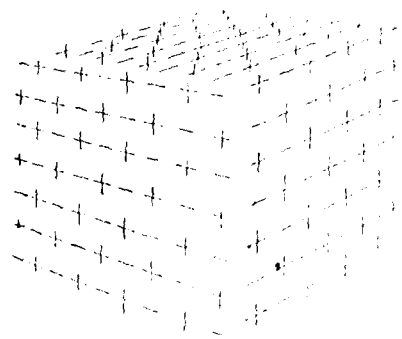


Fig. 17.

It will be noticed that one cannot talk of a sodium chloride molecule, NaCl. In the crystal each Na^+ ion is surrounded by six Cl^- ions and each Cl^- is surrounded by six Na^+ ions (except those in the surface, of which more later). The whole thing is held together by the fact that it is made up of mutually attracting particles.

So much for the ionic or electrovalent link, as it is called. The other way in which atoms can be held together is by sharing electrons. In Figure 18a we see two separate hydrogen atoms. Each consists of a positively charged nucleus with an electron moving round it. The electron spends most of its time in the shaded area which represents a diffuse sphere surrounding the nucleus.

This sphere may be regarded as a cloud of probability, for the probability of finding an electron at any point at any instant is given by the density of shading there. The probability is highest near the nucleus and tails off rapidly as we move away from it. We will regard the sphere as denoting

the region in which this probability is of more than negligible magnitude.

Fig. 18a.

Fig. 18b.

When two hydrogen atoms come together, their electron probability clouds merge, as shown in Figure 18b. We cannot then distinguish between the two electrons, and the density of the cloud at any point gives the probability of finding either of them there. It turns out that the energy of the system is much lower in this state and therefore the arrangement is stable. It is called a covalent link.

In diamond, which is a crystalline form of carbon, each carbon atom may be regarded as at the centre of a tetrahedron with one covalent link extending to each of four other carbon atoms situated at the corners. In the case of covalent links, the electrons are more or less confined between the pair of atoms which they are binding together. They are very strong and highly directional. Hence the hardness of diamond and the rigidity of its crystal.

In metallic crystals, however, the electrons are not confined between individual pairs of atoms, but the probability clouds merge into one another so that, in certain directions at any rate, an electron can travel from one end of the crystal to the other. That is why metals conduct electricity so well. As a crude analogy, it may be said that the system is rather like a fine powder (the particles of which represent atoms) which has been allowed to soak up a liquid (representing the electron probability cloud) so that it sticks together in a compact mass like putty. If we put the mass of putty, say, into a cylinder and apply oil under pressure to one end, it will allow the oil to seep through. This is analogous to the conduction of electrons through a metal.

Until now we have discussed crystals made up of separate atoms. In many substances the atoms are tightly bound together in groups, i.e. molecules. This means that the units of which the crystals are built are not simple spheres, but more complicated in shape. For instance, in benzene the carbon atoms are grouped together in rings of six which may be imagined as bangles, each made up of six large beads. At low enough temperatures these bangles will stick together because, although the electrons spend most of their time in the individual molecules, there are projecting electron probability clouds that can interact weakly to hold the molecules together in a crystal. Most organic crystals are of this type. Because the bonding is weak so are the crystals. A familiar example is naphthalene (or moth balls), whole molecules are made up of double benzene rings fused together like a figure of eight. The weak attractions which give rise to it are called van der Waals forces.

Crystals can therefore be classified not only according to their geometry, but also by the type of bonds which hold them together. We may have ionic, covalent, metallic or Van der Waals bonding, or even hybrid types of link intermediate between two of these.

The Physical Properties of Single Crystals

The physical properties will depend very largely on the chemical bonding in a crystal. Ionic crystals are usually transparent to visible light, and are brittle. When broken, they tend to cleave along certain preferred crystal planes. Rock salt, which has the structure shown in Figure 17, cleaves along the faces with the chessboard pattern and hence it is easy to cleave it into a cube. Not all ionic crystals are brittle. Some can be bent without breaking, though this, of course, destroys the crystal structure. Covalent crystals are hard and brittle, while those with van der Waals bonding are soft and greasy because the separate molecular units can glide past each other easily.

Metal crystals (surprisingly) are soft, but coherent. The

effect of pressure is to cause atoms to flow over each other while remaining in tightly bound contact. This is the way in which our putty in the previous analogy would behave if we squeezed it. This flow is not random, but certain crystal planes tend to slide over each other more easily than others. The use of rafts made up of bubbles to represent two-dimensional crystals has been described in *Science News 1*. These bubble rafts were used to illustrate this effect. Many investigations have been carried out on the stretching of single crystal wires, principally by Professor E. N. da C. Andrade and his fellow-workers.

Because the electrons in a metal are relatively free to move, they can interact with electromagnetic waves in such a way as to reflect them, i.e. metals are good reflectors of light. At high temperatures the electrons can jump right out of the metal (this effect is used in the radio valve) or they can be dragged out at room temperature by powerful electric fields. The ease with which they can do this may depend on the particular crystal face or edge from which they jump.

It may be asked why, if the crystals of which a piece of metal are composed are soft, hard metal objects can be made? The answer is that when we talk of metal single crystals we mean those in which the structure is as perfect and free from strain as possible. But an ordinary piece of metal is made up of a three-dimensional mosaic of crystals, and many of these are under strain because the lattices are distorted to enable them to fit together at the boundaries. Distorted lattice planes will not slide over one another easily and hence polycrystalline metals are often hard.

The Importance of Crystal Surfaces

Almost all solids are made up of crystals, however minute. Provided we mean by crystal a three-dimensional pattern of atoms or molecules, even wood fibres, muscle fibres and other biological products can be regarded as crystalline. Most reactions involving solids will either take

place at crystal surfaces or be controlled by processes so occurring. They include many technologically important processes, such as crystallisation itself, corrosion, and catalysis. The principles involved in these are also important from a fundamental point of view, especially in the case of catalysis, because the atoms in a crystal surface lie in known positions and so we have a good chance of finding out the spatial relationships necessary for the chemical reactions to take place. The value of this can be realised when we consider biochemical catalysts (enzymes). These are very 'choosy' and each will usually catalyse only one reaction. The reason for this specificity is attributed to the fact that each enzyme has a definite electron probability cloud shape (or configuration as it is called), which will interact with only one pair of substances to promote reaction between them. Their structures are so complicated that at present we cannot hope to study them in any detail. Metal surfaces are about the simplest catalysts known.

The Surfaces of Crystals

When we were discussing the atomic bonding in crystals, we did not consider conditions at a surface. Suppose we take a diamond and cleave it so as to expose fresh surfaces. What has happened to the bonds which previously held the two pieces together? Do the electron probability clouds retain their well-defined nature and project from the surface like rows of cabbages, or are the bonds near the surface rather different from those in the interior? How do they behave when other atoms carrying electron probability clouds are brought near? If such atoms become bound to the surface, how strong are the bonds? We would expect the chemical nature of the surface to depend markedly on the crystal plane exposed, for this will decide not only the pattern of the atoms in the surface, but also the angle at which we might expect 'residual valencies' (as they are sometimes called) to emerge.

Ionic crystals can present surfaces containing both positive and negative charges, FG (Figure 19), or only one kind, HI. Again we may expect marked differences in the chemical properties of different crystal faces.

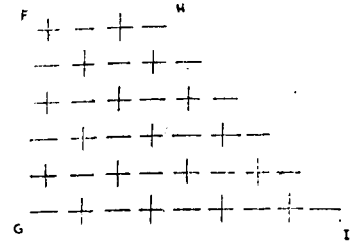


Fig. 19.

Metallic crystals are particularly interesting, since the precise nature of the metallic bond is still under discussion, and also because many metals show marked catalytic properties at their surfaces. Of the substances showing catalytic activity, metals are the most easily made into single crystals of convenient shape and size. For these reasons, most of the rest of this article will be concerned with metal crystals.

The Chemical Reactivity of Crystals

About twenty years ago, Tamman and his collaborators in Germany began growing large single crystals from the melt, and examining their properties. One metal which they studied was copper. Spheres were very carefully turned from single crystal rods and their surfaces freed from 'bruised' crystal material by etching and gentle abrasion until they had a polished appearance. The spheres were then subjected to etching by various acids. These attacked to produce pits and furrows in the surfaces, with their sides parallel to certain crystal planes. Various acids exposed various planes, and even with the same acid the result depended on its concentration. Fig. 20 illustrates the effect for an acid which reveals cubic micro-facets, as they are called. It shows a cross

section taken parallel to one of the cubic planes. The sides of the etch pits are seen in section as steps with sides parallel to the other two cubic planes. The little facets act as mirrors so that in the six directions, OA, OB, OC, OD, and the two perpendicular to the plane of the page, an incident light beam will be reflected back along its own path. This work was of great value in enabling later investigators to find out how the crystal planes lay in any particular specimen of single crystal. The principle of a method which the author has

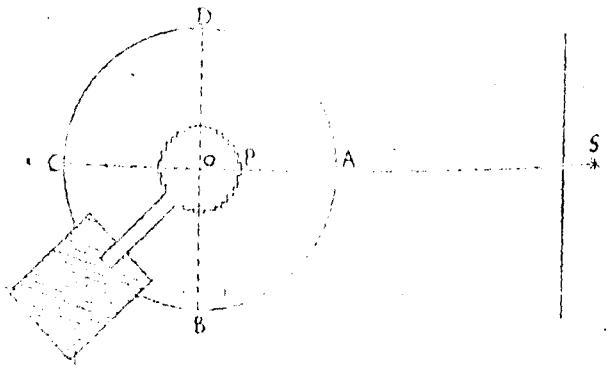


Fig. 20.

found convenient will be outlined. The crystal is etched with an acid which exposes cubic planes and mounted at the centre of a glass sphere (in Fig. 20, the size of the crystal is shown exaggerated). It is illuminated with a parallel beam of light SP and rotated until one set of micro-facets is perpendicular to the beam as shown, so that the reflected beam passes back along the path of the incident beam. The point A at which the light strikes the glass sphere is then marked. Similar marks BCD etc. are made for the other cubic reflections and the directions of the principal crystal planes are thus clearly recorded.

The crystal is then removed, repolished and a reaction carried out on its surface. When it is replaced in its old position in the sphere, the faces affected by the reaction are

easily identified by their positions relative to the spots on the glass sphere.

Tamman's work was followed up by other investigators who studied the oxidation of metal crystals. In particular A. T. Gwathmey and his collaborators in America, working with spherical copper crystals, showed that the behaviour of the faces when the crystal was heated in oxygen gas varied widely. Some were rapidly covered with a thick layer of oxide, while others were hardly attacked at all, so that the surface of the sphere was mapped out in a striking geometrical pattern (Plate 11).

The results are not easy to interpret, for metal oxides are themselves crystalline, and electron diffraction experiments (see later) have shown that the oxide tends to lie so that the crystal lattice fits that of the underlying metal. This in itself may govern rates of oxidation. Furthermore, once a layer of oxide has formed, subsequent reaction must take place by the reactants (particularly the metal) diffusing through it. This process too may be governed by the orientation of the oxide.

Meanwhile much research had been carried out on heterogeneous catalysts, that is, catalysts at whose surfaces reactions between gases are brought about. This led to the view that they consisted of minute crystallites containing only a few thousand atoms each and that during reaction the reagents became attached to their surfaces, probably by the 'residual valencies' mentioned earlier. Theories were also put forward that for a given crystallite certain faces and especially certain edges and corners were more active than others. It was therefore rather surprising that no one thought of trying to carry out catalysis on single crystal surfaces until Gwathmey, in 1940, showed that the reaction between hydrogen and oxygen proceeded catalytically on certain faces of a copper single crystal.

He used an electrolytically polished sphere and passed the hydrogen/oxygen mixture over it at temperatures between 360°C. and 444°C. After about 15 minutes it was found that micro-facets were being formed, i.e. that as the hydro-

gen combined with the oxygen on the surface, its atoms were being rearranged to form crystal planes so that reflections could be obtained from them as from a crystal etched in acid. The experiments were repeated using slices of crystal, exposing only one face at a time, and it was shown that the reaction proceeded more than twice as fast on an octahedral ('cube corner') face as on a cubic face. These experiments are extremely interesting and valuable, for they show not only that reactions can be catalysed by massive crystal faces, but that even at moderate temperatures (compared with the melting point of copper, 1080°C) copper atoms can migrate distances which, relative to their own size, are enormous when a surface is involved in catalysis. They show also that a catalytic reaction of this sort is continually creating its own catalytic surface as it proceeds, and may cause an entirely new outlook on the theory of catalysis in favour of a more dynamic view of the surface atoms.

Gwathmey has similarly shown that only certain surface planes of a nickel crystal will decompose carbon monoxide at high temperatures, that a copper crystal heated in contact with mercury absorbs the latter more on some planes than on others, that when more copper is slowly deposited by electrolysis on a single crystal sphere of copper some faces build up more quickly than others, and so on.

The author has extended the spherical crystal technique to reactions in which no product is built up at the surface. Copper spheres with electrolytically polished surfaces were heated in the vapours of chlorine, bromine and iodine (of the halogen group of elements) to form copper chloride, copper bromide and copper iodide respectively. These substances are volatile and under the conditions of the experiment sublime away as fast as they form. The result is that the reaction removes copper from the surface which becomes etched, i.e. etch pits and furrows are formed with their sides parallel to certain crystal planes (just as when a metal crystal is etched by acid).

Which planes are exposed depends on the vapour used

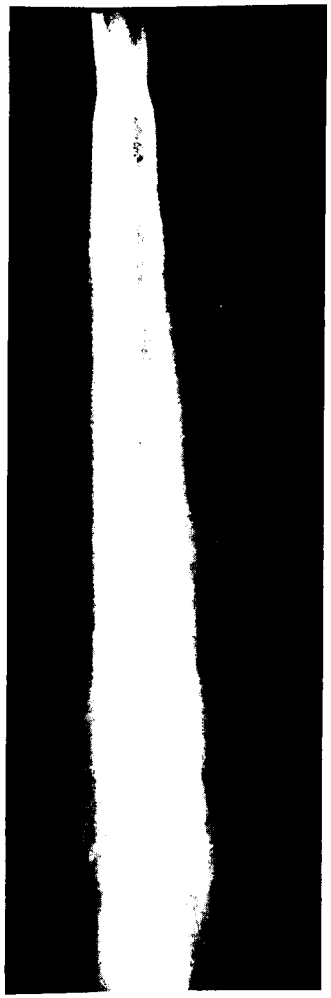


Plate 6. Inner shell of 100 lb. thrust experimental rocket motor, showing the helical cooling passage. (p. 81)

Plate 7. Vickers transonic research aircraft. This model aeroplane, powered by a liquid-fuel rocket and launched from a Mosquito, was designed to investigate the forces acting on an aircraft reaching sonic speed. (p. 83)

(By courtesy of *Flight*)

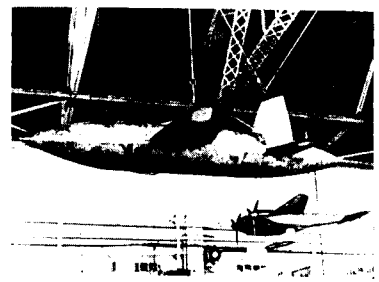


Plate 5. Exhaust from experimental rocket motor. The segmented appearance of the flame is due to the shock wave which is formed in the jet as it leaves the nozzle and which is reflected to and fro within the jet. The abrupt rise of pressure in the shock waves causes the exhaust gases to glow more brightly and thus renders the waves visible. (An explanation of shock waves was given in *Science News* 7.) (p. 75)

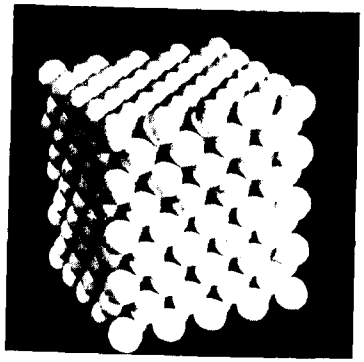


Plate 8. Cubic close packing of spheres cut parallel to planes which show the 'egg crate' pattern. (p. 54)

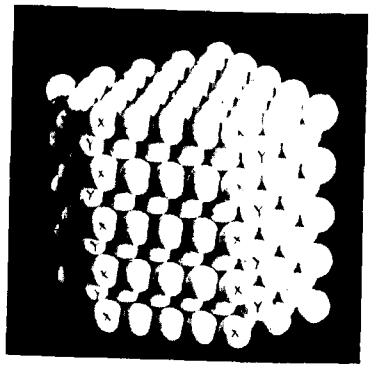


Plate 9. Previous model with edge cut away at 45° to expose a dodecahedral face. Note the different structure with horizontal rows of atoms running 'one high (X – X) one low (Y – Y)' rather like ribbed knitting.

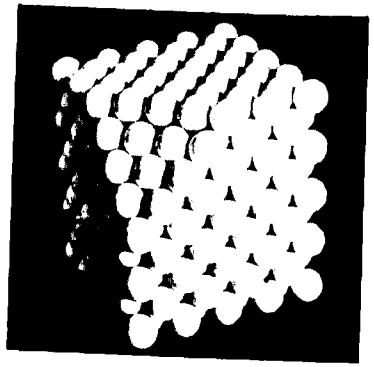


Plate 10. The model of Plate 8 with a corner cut away to expose an octahedral plane. Note that these are the planes of densest packing with the atoms arranged in rows parallel to the sides of an equilateral triangle. Similar planes would be revealed by cutting away any of the other corners of the cube. Any one set of these could have been the original close-packed layers of Plate 8. (p. 54)

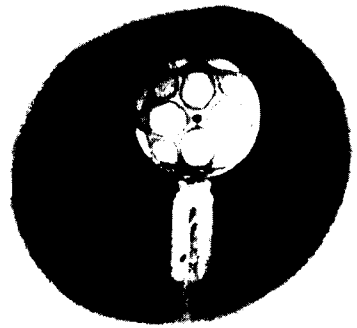


Plate 11. Spherical single crystal of copper oxidised in air at 200 C for ten minutes by Gwathmey and Benton. Different surface crystal planes have oxidised to different extents. The black spot is at the centre of a 'cube face' area. (p. 65)

Plate 12. Micrograph of octahedral ('cube corner') area of spherical copper crystal which has been etched with iodine vapour at 570 C (magnification × 80). The dark areas are the bottoms of pits, the sides of which rise in a series of terraces. (p. 67)

SOME TECHNIQUES OF METALLURGICAL INVESTIGATION

Plate 13. The pattern of flow in a model Bessemer converter with side blast, showing the flow of 'gases' above the 'steel' surface in a given plane of the model. Air bubbles are used here to make the flow of water visible; water is used to simulate the air blast of the real converter. (p. 141)

(By courtesy of the Iron and Steel Institute.)

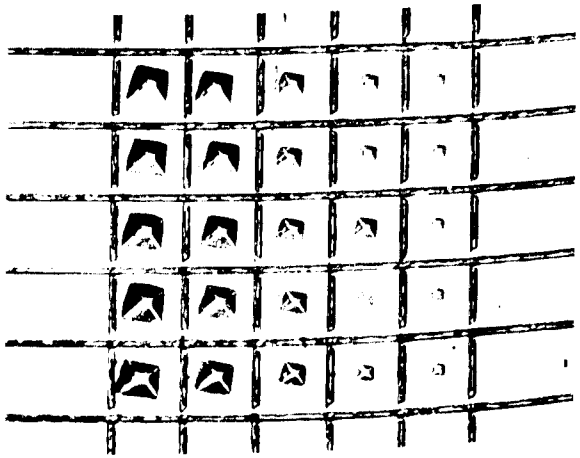


Plate 14. Squares of 0.02 mm. side ruled on Austenitic stainless steel to indicate the precision with which the impression can be made. The loads were 50, 40, 20, 10 and 5 gm. (× 475). (p. 150)

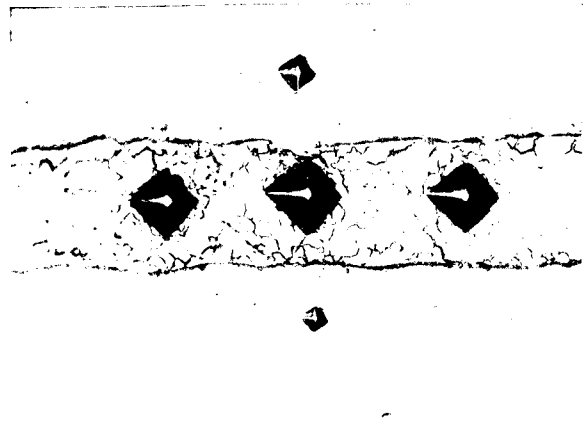


Plate 15. Tin foil, 0.05 mm. thick, between plates of brass and of gun-metal. Load 5 gm. (× 475). (p. 150)

Plate 16. Slip-lines in pure zinc after exposure to 50 thermal cycles between 30°C and 150°C (× 500). (p. 151)

(Plates 14 to 16 by courtesy of the Institute of Metals).

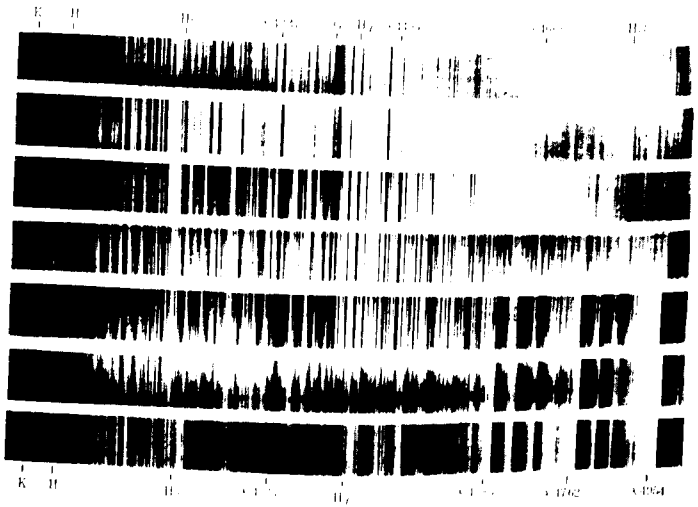


Plate 17. Typical Stellar Spectra of Classes G, K, M
(By courtesy of The Observatory, University of Michigan, Ann Arbor)
(see page 44)

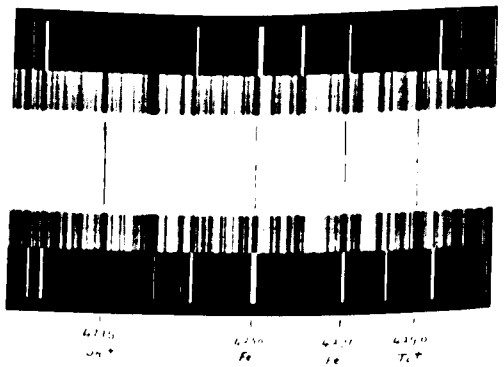


Plate 18. Spectrum of β Leporis and α Centauri
(By courtesy of the Lick Observatory, U.S.A.)
The centre strip carries ink lines to show the lines of ionised strontium (Sr $^+$) and of iron (Fe) used for measurement.
(see page 44)

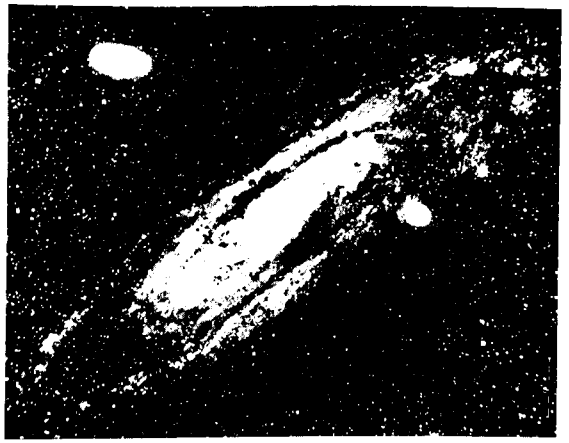


Plate 19. The Nebula in Andromeda.
(By courtesy of the Yerkes Observatory, Williams Bay, U.S.A.)
(page 47)

Plate 20. Small Magellanic Cloud.
(By courtesy of the Harvard College Observatory, U.S.A)
(page 46)

BLOOD GROUPS

Plate 21. A simple blood - grouping test. The anti-A serum on the left has had no effect on a drop of the patient's blood. The anti-B serum on the right has agglutinated the blood, and the patient, therefore, belongs to Group B. (Natural size). (p. 109)

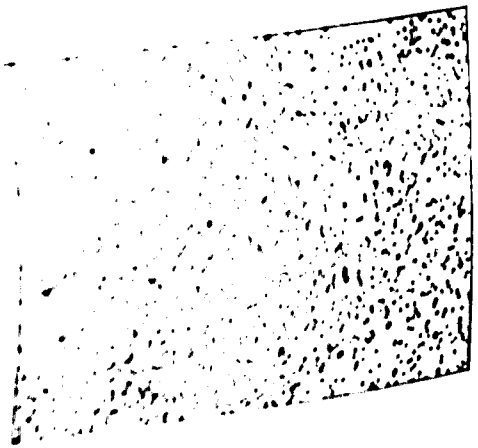


Plate 22. The appearance of normal red blood cells when viewed under the microscope. (Magnification × 75).

Plate 23. The appearance of agglutinated red blood cells under the microscope. It can be seen that the cells adhere together in large clumps. (Magnification × 75). (p. 105)

and on the temperature at which reaction takes place, but a given set of conditions always exposes the same set of micro-facets. For instance, iodine at 570°C always reacts to expose octahedral planes. Plate 12 shows part of the octahedral face of such a sphere magnified 80 times. The etch pits are seen to go down in a series of terraces like Dante's Inferno. The upper surfaces of the terraces are octahedral micro-facets and the sides also are probably made up of such facets parallel to others of the four octahedral ('cube corner') planes.

This type of reaction has many advantages. There is no layer of solid at the surface to complicate the reaction. The micro-facets produced can be identified accurately by light reflection as indicated. Once the early stages of the reaction have destroyed the original polished surface, the facets revealed are those at which reaction takes place. The layers of atoms must be stripped off them one after another so as continually to reveal fresh layers with the same atomic pattern. The geometry of these patterns is known in detail and we have therefore a comparatively simple system to study, i.e. reaction between a vapour and metal atoms in a fixed pattern. The bearing of this on problems of residual valency etc. is fairly obvious.

Another way in which surfaces may be studied is by electron diffraction. Crudely speaking, a beam of electrons is 'bounced off' the surface (in a vacuum) and collected on a photographic plate. Such beams are simply scattered by non-crystalline material, but, because they have wave properties, they are reflected like light in a regular way from crystals, giving a pattern of spots on the plate. These patterns are related to those of the atoms in the surface and can be interpreted mathematically, giving the orientation of the crystal surface and the spacing of its atoms. The electrons do not penetrate very deeply, so if a film of another substance is formed on the surface we obtain the pattern belonging to that substance. Thus we may take a crystal surface and, having determined its nature by electron diffraction, deposit another substance on it as a thin film from the vapour, or by

oxidising the surface, or by electrolysis etc., and then re-examine it. The film itself may be crystalline, and if so it is often found that it will be built up in such a way that the two atomic patterns, that of the film and that of the underlying crystal, fit into each other at the boundary. This may require some stretching or compression of valency bonds so that the boundary between the two is under strain.

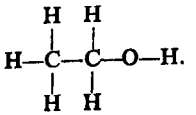
By this means a lot of information has been obtained about films on surfaces which is of very great importance in problems of corrosion (and its reverse, protection), resistance to wear, light reflectivity, electron emissivity etc., and again in the problem of the residual valencies. Much of this work has been done by Professor Sir G. P. Thomson, Professor G. I. Finch, and their colleagues in Britain.

What happens at a surface may be determined by what is going on under it as well as over it. Consider what happens when a clean surface, e.g. of metal, is exposed to a gas. It has been shown that even a liquid mercury surface rapidly collects a film one molecule thick, a so-called adsorbed layer. In some cases adsorption is very marked, even with gases which are not very reactive at the temperatures of the experiment. Finely-divided palladium will adsorb enormous quantities of hydrogen and theories differ as to whether the latter is truly adsorbed on the large surface presented by the crystallites or whether it actually passes into solution in the metal. Some light may be thrown on the situation by the fact that a thin membrane of palladium metal at an elevated temperature will allow hydrogen to diffuse through more readily than will other metals. W. R. Ham and J. H. Simonds in America have developed a theory that gases diffuse easily through metals only when they react with them, i.e. when the electron probability clouds of the gas atoms can merge with those of the metal. A hydrogen atom may thus loose its electron which becomes just 'one of the crowd' holding the crystal together, while the positive nucleus diffuses into the chinks between the metal atoms.

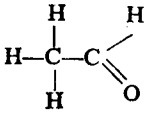
The question as to whether diffusion takes place along the

boundaries between the crystal grain in a polycrystalline metal or through the crystal lattices themselves has been answered by Smithells and Ransley in England. They prepared a thin tube of pure iron with a closed end and annealed it until it consisted of only two large crystals. At high temperatures this tube would permit hydrogen to diffuse out at a rate well above that which could be accounted for by diffusion through the single grain boundary. It must have gone slap through the crystal lattices.

The importance of this to the science of catalysis may be seen if we consider a reaction in which a compound has hydrogen removed from its molecules. Such dehydrogenation reactions frequently occur at catalytic surfaces. In the example shown, ethyl alcohol (I) is converted to acetaldehyde (II) by the loss of two hydrogen atoms. If the hydrogen



I



II

atoms can travel through and under the catalyst surface as well as over it, the number of possible explanations of the mechanisms of such a reaction is greatly increased. One view may be that the surface acts as a sort of cabbage strainer, holding back the larger atoms and molecules, but permitting those hydrogen atoms which are not too tightly bound to the parent molecule to migrate away. It is a fact that metals which allow hydrogen to diffuse fairly easily are also good catalysts for this type of reaction.

There is ample evidence that at elevated temperatures the atoms in crystals are by no means fixed in position. A two-dimensional analogy may be made by considering a draught board on which each draught must always sit on a square but not necessarily on the same square all the time. The draughts can hop from square to square. An experimental example of this self-diffusion (as it is called) may be cited in the work of Banks, Day and Miller in America. They elec-

troplated the end surfaces of single crystal zinc rods with a thin layer of radio-active zinc, heated the rods to just below the melting point of zinc for several hours, then took samples at varying depths from the end surface. By measuring the radio-activity of the samples, it was possible to estimate the rate at which the radio-active zinc had diffused. They showed that the way in which the shuffling of the zinc atoms took place between one lattice layer and the next was quite definitely dependent on which particular set of lattice planes was concerned.

If the draught board is completely filled with draughts (i.e. there are 64), it is difficult to see how the draughts can shuffle unless they change places by all jumping at once, and this is not very probable.

But if there are vacant sites this difficulty is obviated, for draughts from adjacent sites can jump on to them, leaving their sites vacant, and so on. Now theory indicates that even a perfect crystal will contain a certain number of vacant lattice positions. The proportion of such sites depends only on the temperature and not on the degree of perfection of the crystal, and it is believed that self-diffusion within a crystal is due to the shuffling which such gaps make possible. Substances may be able to diffuse through a crystal in two ways, either through the chinks between atoms, or by replacing atoms and taking part in the shuffling. Experiments on diffusion enable us to measure these effects and possibly to relate them to surface phenomena.

To sum up, a crystal can be classified according to two sets of considerations:

(*a*) The type of bond which holds it together (electrovalent, covalent, van der Waals, metallic).

(*b*) The three-dimensional geometrical pattern into which the units (atoms or molecules) of which it is composed, fit.

Consideration of the chemical reactivity of crystals will obviously be focussed primarily on the surfaces. Even in the same crystal the surfaces may vary enormously in physico-chemical properties. Investigation of these properties is

important to the whole study of chemical reactivity. Reactivity may also be influenced by diffusion within the crystal.

This article does not pretend to have covered the whole field. It is intended simply to indicate that a new branch of science is developing to explore and exploit the possibilities outlined above. It is the study of crystal reactivity.

My thanks are due to Mr K. J. Fereday for the diagrams.

Original papers in which further references will be found are:

(a) A. T. Gwathmey and A. F. Benton, *Journal of Physical Chemistry*, **44**, page 35, (1940).

(b) and **46**, p. 869 (1942).

(c) A. T. Gwathmey and H. Leidheiser, *Journal of the American Chem. Society*, **70**, p. 1200 (1948).

(d) G. Tamman and F. Sartorius, *Zeitschrift für anorganische Chemie*, **175**, p. 97 (1928).

(e) *Some Problems in Adsorption* by J. K. Roberts, Cambridge Univ. Press.

(f) The literature on electron diffraction studies is rather large, but a start may be made by looking up the work of G. I. Finch in the latest decennial index of the *Chemical Abstracts* of the American Chemical Society.

Rockets

D. HURDEN

A ROCKET is a jet propulsion device which burns a fuel of some sort and expels the gaseous combustion products rapidly backwards. So far the definition applies equally to the aircraft jet engine which is by now familiar to most people, but there is one important difference. In order to burn any fuel a supply of oxygen must be available. The ordinary jet engine uses oxygen from the air, but the rocket carries its own oxygen supply with it in some form. It is thus completely self-sufficient and would work just as well outside the Earth's atmosphere where there is no free oxygen as in it.

This immediately raises a point which is a stumbling-block to a lot of people, not excluding some engineers who ought to know better. Since news of Whittle's jet engine was released there has been a popular misconception that if there is no atmosphere for the exhaust gases to push against there will be no thrust; that, although a rocket may burn in a vacuum, it will not be able to push itself along. But this is not true. A rocket is another demonstration of a famous law first stated by Isaac Newton in 1687 which says that 'action and reaction are equal and opposite', meaning that if you exert a force on some object, for instance throw a cricket ball, the cricket ball will exert an equal force on you in the opposite direction, and if you throw the ball hard enough, this force will make you overbalance backwards. Another familiar example is provided by a fire brigade. Those who have watched fire-fighters in action will know that the nozzle is by no means easy to hold, that, in fact, the reaction produced by the jet of water is often as much as two men can stand. Neither of these ways of producing thrust needs the

surrounding air to do it, nor does a rocket, which produces its thrust by squirting out large quantities of gases in the form of a high speed jet. I have dwelt on this point because the fallacy seems to be very deep-rooted, but if some people are still unconvinced, I will add that the theory was confirmed in America by Dr Robert H. Goddard who, by 1919, had fired small rockets inside an evacuated tank and measured the thrust exerted by them on their anchorages.

From the fact that rockets in a vacuum not only burn but will continue to exert a thrust, it follows that some sort of rocket is the only practicable means of propulsion for travelling to other planets.

The rocket is not a new-fangled idea. The Chinese are alleged to have had rockets as long ago as the 13th century, but it is not until the end of the 18th century, when they were used by the British Army in India, that we have positive historical evidence of their use. When their spell of usefulness as weapons was temporarily over they remained a valuable asset to coastguard services, and line-carrying rockets fired to wrecks have saved hundreds of lives. The design of these had not changed much by the beginning of this century, when Goddard's name crops up once more in the story. He measured the efficiency of an ordinary ship rocket and found it to be only 2 per cent, a surprisingly low figure when you consider that the corresponding efficiency of an aircraft jet engine is about 25 per cent. By simple alterations he succeeded in raising the efficiency of his rockets to 30 or 40 per cent, but before we can appreciate his achievement fully we must examine a little more closely the behaviour of the gases leaving a rocket.

Fig. 21.

Fig. 21 represents a rocket in which some kind of fuel or, more exactly, propellant is burnt. The gases produced escape from a small hole in one end which acts as a vent; the smaller the hole, the greater will be the restriction and the higher will be the pressure generated inside the casing.

Now when a gas under pressure escapes from a plain hole like this, it can be shown that its pressure falls until it reaches about half its original value as the gas actually flows through the hole. When it leaves the hole at this pressure it undergoes a sudden expansion down to the pressure reigning outside. Thus the gases moving away from the rocket along its axis (which are usefully employed in propelling the missile) contain only part of the energy initially possessed by the combustion gases, in the form of velocity acquired as they expand down to half their original pressure, while the extra velocity that they might have acquired if they had been allowed to expand gradually down to atmospheric pressure is lost by this sudden random expansion.

Suppose that we shape the entrance to the hole and follow

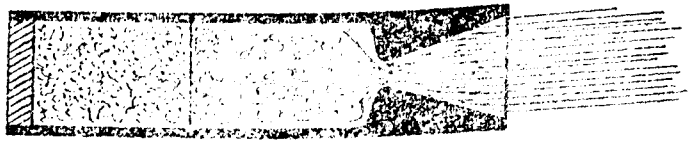


Fig. 22.

it by a passage of gradually increasing diameter as shown in Fig. 22. Then the gases escaping through the hole will not be able to expand immediately, but because the passage increases in area will expand gradually as they travel towards the end, so that by the time they have reached the outlet they will be moving at a considerable speed away from the rocket. This device transfers the energy in the high-pressure gases to a fast-moving gas stream leaving the rocket, and if the passage is properly designed the gases will leave the nozzle at the same pressure as the surrounding atmosphere.

Returning for a moment to the cricket-ball analogy, we know that the faster we throw a ball the more likely we are

to fall over; so in the same way a rocket burning fuel at a given rate will give more thrust the faster the gases leave it. Of course, it will also give more thrust if it burns fuel at a greater rate, but we will ignore this complication. We have already seen that we can increase the speed of a jet of gases leaving a rocket of a given size by adding a divergent passage to it. This was first realized by Goddard in 1916, although it was common practice in steam-turbine work by 1890, and it was this simple addition, together with a considerable increase in working pressure, made possible by the use of a high tensile steel casing, that increased the efficiency of Goddard's rockets from 2% to over 40% and revealed fresh possibilities in this means of propulsion.

So far we have not said much about the fuels we can use. First of all, there is no restriction on state; a rocket will burn gases, liquids, or solids. Gaseous fuel rockets are not much use in practice, since to carry the necessary weights of fuel and oxygen in the form of gases would need far more space than usually exists in an aircraft or missile. However, they are commonly used for research work, since the necessary quantities of gas can be stored on the ground, and rockets burning hydrogen and oxygen or methane and oxygen have been used for hundreds of tests (Plate 5).

Solid fuel rockets are the oldest variety. In China the propellant was gunpowder, while Goddard's experiments were done with various smokeless powders. To-day solid fuel rockets are used mostly as medium-range military or anti-aircraft weapons or else to give extra thrust for a few seconds to a heavily-loaded aircraft taking off. During World War II the British developed 2″, 3″ and 5″ calibre rockets which burned cordite and supplemented very effectively our orthodox artillery. In America JATO ('jet-assisted take-off') rockets fitted to aircraft have been made to burn all sorts of things, not usually thought of as propellants, such as asphalt.

Most work, however, is being done in the field of liquid propellants, which is also probably the most recent develop-

ment, dating as it does from the early 1920's. The more obvious liquid fuels such as petrol were tried first, and the oxygen needed to burn them was supplied by feeding liquid oxygen into the rocket at the same time. Liquid oxygen boils at −183°C at atmospheric pressure and will soon evaporate if left to itself in an open vessel, so it has obvious disadvantages as a rocket propellant; nevertheless it always has been and still is very popular. Since the early days literally hundreds of possible fuels and oxygen-carriers have been tried, but before amplifying this bold statement let us see what conditions rocket propellants must satisfy.

There are two main requirements: first, the combination must be theoretically suitable, and, second, it must not be too difficult to handle. We have seen already that the faster the jet of gases leaves the rocket the greater will be the resulting thrust, and it is not very difficult to prove that the velocity of a jet of gases expanding through a divergent nozzle from a given initial pressure varies as the square root of their absolute temperature, and inversely as the square root of their molecular weight. We can write this as

$$v \propto \sqrt{\frac{T}{M}}$$

Now the most useful measure of the efficiency of a rocket is called 'specific impulse', a most misleading term since it is not an impulse at all but the thrust produced by burning one pound of propellant per second. Thus

$$I_s = \frac{\text{Total thrust}}{\text{Total mass flow}}$$

and by using another of Newton's laws of motion we can relate this to the jet velocity. A force is the product of a mass and an acceleration, so if W pounds of gas are accelerated to a velocity of v feet per second in one second, they will exert a force of Wv poundals, or $\frac{Wv}{g}$ pounds. The thrust of a rocket burning W pounds of propellants per second and

expelling the gases at v feet per second will therefore be $\frac{Wv}{g}$ pounds, and the specific impulse will simply be equal to $\frac{v}{g}$. Hence the specific impulse will also vary as $\sqrt{\frac{T}{M}}$. It is possible to work out the temperature attained in a chemical reaction between two propellants, also what products will be formed and hence their average molecular weight, so that we can predict beforehand what specific impulse we can expect from a rocket using them. A few typical figures are given in the table below.

Table 1

Propellants	*Specific Impulse*
Aluminium + oxygen	560
Magnesium + oxygen	560
Hydrogen + oxygen	350
Petrol + oxygen	240
Ethyl alcohol + oxygen	240
Aniline + Nitric Acid	220
'C - stoff' + 'T - stoff'*	210

Besides these theoretical considerations there are many practical points that decide the value of propellants. For instance liquid oxygen and liquid hydrogen have a good specific impulse, but, as they are both liquids with very low boiling-points, they create troublesome storage problems. Nitric acid and hydrogen peroxide are both rich in oxygen and are both liquids at normal temperatures, but they are also extremely corrosive and unpleasant to handle.

The choice may also depend on the way in which the propellants behave when they are mixed. Propellants can be divided into several classes according to their reactions with

*German code names; C - stoff was a mixture of 57% methyl alcohol, 30% hydrazine hydrate, and 13% water, while T - stoff was simply hydrogen peroxide of 87% concentration.

each other. One class, called hypergols, react as soon as they come together, a property which eliminates the need for an igniter of some sort but which calls for great care in handling the propellants. Such pairs are aniline and nitric acid or hydrazine hydrate and hydrogen peroxide. A variation of the second system was used in the Me 163, employed as an interceptor fighter in World War II in which the hydrazine hydrate was mixed with methyl alcohol and water. The group possessing opposite properties to the hypergols, called simply nonhypergols, includes the pairs such as petrol or alcohol with liquid oxygen which do not ignite on mixing. A third group, which is becoming of increasing importance, is the monergols, which contain the oxygen needed for their own combustion. Members of this group, such as nitromethane (CH_3NO_2), tend to be unstable, and research is now being intensively carried out with the object of making them safe to handle, as they undoubtedly simplify matters by eliminating the separate oxidiser.

Work on fuels is one of the fundamental lines of research now being carried out. Hundreds of possibilities have to be explored. Fuels such as furfuryl alcohol and orthotoluidine, which have been little known outside the laboratory, may be just what the rocket engineer wants. Table I shows the possibilities of the light metals aluminium and magnesium as fuels, but research into the best way of burning them is needed. One way that has been suggested is to prepare a fine suspension of them in oil, which can be injected in the ordinary way. Again, theory suggests that liquid fluorine could be used as an 'oxidiser' with suitable fuels, but great care is needed when handling this liquid, as it is very reactive, especially with organic materials. Yet again, it has been suggested that liquid ozone should be used as an oxidiser, since it contains even more oxygen in a given volume than liquid oxygen itself, but although ozone can be liquefied it is very unstable and tends to decompose explosively. Here, then, are three of the problems facing rocket scientists, and there are many more in the field of fuels alone.

Now that we have seen a little of what can be done it is very natural to ask how it is done. How are all these principles applied?

A liquid rocket motor is the most difficult example, so let us confine our attention to that. Obviously we must have tanks in which to store the propellants and some kind of 'combustion chamber' in which they can be burnt. Also we must have some means of transferring the fluids from their tanks to the combustion chamber, and we shall probably want valves to control their flow. We will follow through the system and mention a few of the problems to be faced at each stage.

The tanks are relatively easy, but various propellants have their own special requirements. Liquid oxygen tanks can be made of almost any metal, but they have to be lagged to reduce the loss of oxygen by boiling. On the other hand hydrogen peroxide is decomposed into steam and oxygen by almost every metal except very pure aluminium, but remains liquid at normal temperatures.

There are two main methods of feeding the propellants into the combustion chamber and these are shown diagrammatically in Figs. 23 and 24.

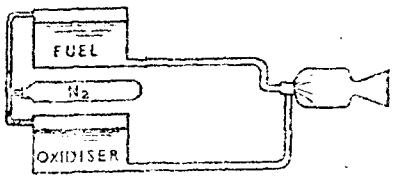


Fig. 23.

In the pressurised tank system shown in Fig. 23 the liquids are displaced from the tanks by an inert gas. Nitrogen, stored under a high pressure and fed to the tanks through pressure-reducing valves, is commonly used for this purpose.

If a rocket has to run for longer than 30 seconds or so, a prohibitive amount of nitrogen would have to be carried to displace all the fuel, and in such a case pumps would be used.

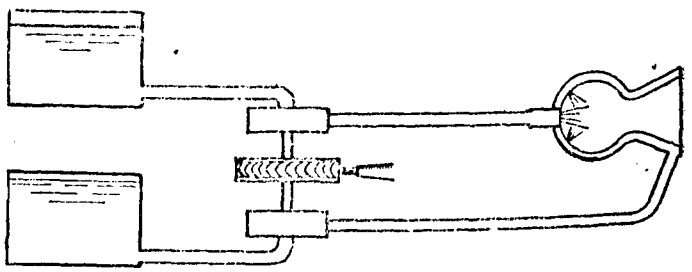


Fig. 24.

The V2 rocket, shown diagrammatically in Fig. 24 (and in real life in Plates 1–3, *Science News 12*), ran for about a minute, and alcohol and liquid oxygen were fed into the combustion chamber by pumps which were driven by a simple steam turbine, the steam being obtained by the chemical decomposition of hydrogen peroxide. The design of rocket pumps again demands careful choice of materials, while pumps for low-temperature liquids require a great deal of research before bearings and sealing glands can be found that are still effective under conditions of extreme cold.

The third major item is the combustion chamber, which may include some means of lighting the propellants and certainly some means of injecting them; this chamber is the limiting factor in rocket design at the moment.

Before the propellants can be burnt, they must be injected into the chamber in a suitable form. Usually a fine mist consisting of a mixture of the two is aimed at, since, in general, the finer the droplets and the better they are mixed, the faster they will burn. There are scores of ways of doing this; the choice will depend principally on the propellants themselves and the shape of the combustion chamber.

The propellants being mixed, they have to be lighted, and again there are several ways of doing it. Some propellants ignite themselves when they mix, otherwise a flame from some kind of cartridge or (as in the case of the V2) fireworks will do. Another way which has been suggested is to inject a shot of some fluid such as zinc diethyl which ignites spontaneously as soon as it meets the air, so that when the fuel and oxidiser arrive there is already a flame to ignite them.

The combustion chamber itself has the hard task of containing gases at a temperature of more than 2,000°C, and usually at a pressure some 20 or 30 times that of the atmosphere. The pressure decides that the chamber must be made of metal for reasons of strength, but all the common metals and most of the uncommon ones melt at temperatures far below 2,000°C, so that if it is to last for more than a few seconds the chamber must be cooled somehow. The cylinders in a motor-car engine present a similar but not such an extreme problem, which is solved by surrounding them with water which absorbs and carries away the heat. Rocket motors that run for more than a few seconds can be cooled in much the same way by enclosing them in a jacket through which a cooling fluid flows. In research motors this may be water, but a much neater way which avoids the use of another liquid is shown diagrammatically in Fig. 24. You will see that in this system the fuel flows through the cooling jacket before being fed into the chamber. This is called regenerative cooling, and it has the added theoretical advantage that some of the heat lost to the cooling fluid is returned to the combustion chamber to do useful work, so increasing the efficiency of the motor.

It may be found that the rocket overheats in spite of this cooling, especially at the throat where the temperature and velocity of the gases are both high, but matters can be improved by speeding up the fuel as it flows through the jacket, since, broadly speaking, the faster it scrubs the surface of the chamber the more heat it will pick up. A simple way of doing this is to make the liquid travel further in the same time by forcing it to take a corkscrew path. The inner shell of such a combustion chamber is shown in Plate 6; this is normally surrounded by a closely-fitting jacket.

We have seen that the efficiency of a rocket varies as $\sqrt{T}$, so efforts are continually being made to run at higher

temperatures. Additional cooling is needed if chambers are to stand these increases, and two possible methods can be mentioned. One of these, known as film cooling, is illustrated in Fig. 25.

Fig. 25.

Small holes are drilled through the wall of the combustion chamber so that fuel can leak through from the jacket. Before it has a chance to mix with the hot gases it boils and forms a protective film of comparatively cool vapour between these gases and the metal wall. The other method is an extension of this idea. It is called sweat cooling and has been made possible by the development of porous metals. The combustion chamber is made of spongy material through which the cooling fluid seeps. On its way through it picks up some of the heat flowing outwards, thus reducing the net heat flow through the wall, and as in film cooling it evaporates when it reaches the inner surface and forms a cool film of vapour.

No mention has been made of the importance of the shape of the combustion chamber, but enough has been said to show that there is plenty of scope for research on this alone.

Now what is the use of all this research? The obvious use of a rocket is as a weapon, either offensive or defensive. Attempts have also been made in the past to use rockets for carrying mail, and although none of these met with any great success, modern methods of control make this quite a feasible proposition. Letters from Europe could reach America in an hour. Another commercial use is the assisting of heavily-loaded aircraft to take off. This is not new; rocket assisted take-offs by heavy bombers are quite common, but a less familiar application would be the reduction of the take-off run needed by a large flying boat.

In addition to their possible commercial and strategic value, rockets are powerful research tools both in aerodynamics, with which they may be closely concerned, and in physics. Their value in aerodynamics lies in their ability to propel full-sized or model aircraft (perhaps remotely controlled) at supersonic speeds in order that the unknown aerodynamic forces in play at these speeds can be measured. An example is the Vickers transonic research model (Plate 7) which was launched in 1947 from a Mosquito and which was intended to transmit to the ground details of the forces acting on its structure as the rocket unit accelerated it up to sonic speed. Similar information was obtained in America by Northrops who mounted wing sections on trucks which were accelerated along a 2,000 foot length of railway track up to a speed of 1,000 m.p.h.

Physicists in the United States have lately collected many fresh data. Until the Americans built a rocket of their own design, most of the captured V2's launched in New Mexico were packed with instruments designed to tell scientists more about our upper atmosphere. The information collected by these instruments was either transmitted to the ground automatically or the instruments themselves were fitted with little parachutes and then thrown out when the rocket began its downward career. In this way we have learned a great deal about the pressures and temperatures about 100 miles above the Earth's surface, and the strength of the Earth's magnetic field, and also added to our store of facts about cosmic radiation. In addition to the sending up of Geiger counters, the effect of cosmic radiation on living creatures was investigated by including a box of fruit flies in the cargo of one rocket. The flies landed by parachute, apparently none the worse!

The measurements of cosmic radiation and magnetic field may be very important in one as-yet-untried application of rockets – interplanetary flight. It is essential to know what

happens to living creatures exposed to these rays, while the Earth's magnetic field may effect the navigation of future space-pilots.

This idea is not as fantastic as it sounds. Many engineers and scientists are seriously considering the ways and means of realizing it before this century ends. The difficulties are great – the first one being to leave the Earth at all. Everybody knows that a cricket ball thrown vertically upwards will come down sooner or later under the influence of gravity. In order to leave the Earth completely, a rocket must do work against gravity, and it can do this in two ways. It can produce a continuous thrust that will keep it moving vertically upwards at a speed of say 500 m.p.h., and in due course can reach the Moon or wherever it has to go, having consumed an amount of fuel that stuns the imagination. The initial size of such a rocket would be fantastically great. On the other hand, if an object leaves the Earth with a high enough speed it will never come back; it will keep moving away with ever decreasing speed. The lowest initial speed for this to happen is about 7 miles a second, so if a rocket could be accelerated to this speed in a reasonably short time, it would then coast the rest of the way to its destination.

This idea has provoked much argument and discussion. The jet velocities obtainable from present-day chemical fuels are of the order of 1 mile a second, while the best velocity theoretically obtainable is only of the order of three miles a second. However, without going into details, interplanetary flight is possible though complicated when chemical fuels are used.*

A neater way of reaching the 'escape velocity' is to use nuclear power, and no doubt research will sooner or later be directed towards this end. Even this solution is not altogether straightforward, however. Although atomic nuclei

*By launching small rockets from larger ones already in motion until finally a rocket is launched at the escape velocity. A two-step rocket has already reached an altitude of 250 miles.

are ejected from their source at very high velocities, perhaps 5,000 miles a second, they leave in all directions and would somehow have to be directed to form a jet. But besides this, an enormous amount of energy would have to be dissipated in forming a propulsive jet of atomic nuclei.

A more promising method uses the principle adopted now for producing industrial power from atomic energy by using it to heat up some working fluid which is then used in an orthodox way. In the case of a rocket, the nuclear reaction would heat up the working fluid (probably liquid hydrogen) to any desired temperature, and the fluid would then be expanded through a nozzle in the usual way. Nuclear energy allows a working fluid with a conveniently low molecular weight to be used and heats it to a higher temperature than any chemical reaction could produce.

Although there will be scope for research on chemical rockets for several years yet, it seems likely that atomic energy will play an increasingly large part in rocket development of the future.

The Blood of Water Fleas

PROFESSOR H. MUNRO FOX, F.R.S.

THE title of this article suggests a subject both trivial and specialized. I shall show, as I go along, that it is neither. Its scope is in reality very wide, and we shall see how researches that are apparently narrow may have the broadest implications. Recent studies of the blood of water fleas have turned out to be linked closely to the physiology of ourselves and of other animals.

First of all: what are water fleas, and do they really have blood? Water fleas are not fleas, but they do live in water. Enlarged colour photographs of them are shown in Plates 1 to 3. Their size is about that of a flea, but they are crustaceans, not insects like the flea. That is, they are relations of the lobster, crab and shrimp. Their scientific name is *Daphnia* and they live in ponds and lakes. They swim continuously by oar-like movements of the big limbs (antennæ) seen in the photographs, and their jumping gait has given them the name of water fleas. By this name they have been known for centuries.

Most of the Crustacea, including lobsters, crabs and shrimps, have a heart. In the transparent water flea it can easily be seen with a microscope, beating rapidly, at a rate between once and twice a second. The heart propels a fluid which has cells floating in it, and these corpuscles can be seen moving round the body. This moving fluid is the blood of *Daphnia*. It is not confined to blood vessels as our blood is; it moves through spaces between muscles, nerves and gut, but it is none the less blood.

We are used to blood that is coloured red. This colour is due to a pigment named hæmoglobin, which from the chemical point of view is a protein. It colours the so-called

red corpuscles of our blood. Hæmoglobin is called a respiratory blood pigment. This is because of its capacity of attaching oxygen to itself and easily giving up the oxygen again where there is little oxygen in the surroundings. When oxygen is attached to the pigment we call it oxyhæmoglobin. In the laboratory a solution of oxyhæmoglobin gives up its oxygen to a vacuum. Thanks to the hæmoglobin in them, the red blood corpuscles of our blood carry oxygen from lungs to muscles and brain, which all the time use it up. The considerable amount of hæmoglobin in our blood enables it to carry about forty times as much oxygen as the blood could ferry without its red corpuscles, and sometimes we need all this oxygen.

We are used to blood that is red, but some lower animals have green blood, others blue, and yet others have colourless blood. The blood of water fleas can be red or colourless, and this is the main theme of the present article. But before coming on to it, I want to speak of a blue pigment found in the blood of other crustaceans, namely the lobsters and the crabs. These animals have in their blood a respiratory pigment which is much less familiar than red hæmoglobin. Indeed its existence is unknown to most people. It is a blue protein called hæmocyanin. This pigment carries oxygen from the gills of the crustaceans to the various parts of their bodies, doing the same duty as hæmoglobin does for us. The only other animals with the blue hæmocyanin in their blood are scorpions, some spiders and certain molluscs. In the last-mentioned group of animals, the octopus has blue blood and so has the snail.

The blood of the octopus, of the snail and that of some prawns is blue, but that of crabs, although it contains hæmocyanin, cannot be called blue blood, because the colour here is masked by a yellowish pigment also present in the blood. Chemically this is a carotenoid-protein. It is so called as it consists of a protein united to a pigment which is nearly related to the colouring matter of carrots. Chloroform, shaken with crab's blood in a test tube, breaks the chemical

link between the two components of the pigment and then dissolves the yellow carotenoid. Therefore, shaking crab's blood with chloroform reveals its blue hæmocyanin; this blue pigment is untouched by the chloroform, which settles out in a separate layer in the test tube, taking the yellow carotenoid with it.

I will say very little more about hæmocyanin, partly because relatively little is known about it, and not much has been discovered recently, but mainly because I want to get on to talk of new discoveries about hæmoglobin in water fleas and other crustaceans. Before leaving blue hæmocyanin, however, it is worth mentioning two interesting points about its molecule. One is that this molecule is the largest of any known substance, having a molecular weight of several millions. The other is that it contains copper. This copper corresponds to iron which is present in the hæmoglobin molecule.

One is perhaps surprised that a snail or an octopus should be able to find enough copper in the food or in sea water to build up the blood pigment. But animals have an astonishing capacity for accumulating elements that are rare in their surroundings. Some marine animals, for example, collect vanadium, an element that is not detectable in sea water. One is also a little surprised to find a copper compound in an animal, when copper salts, such as the sulphate, are so toxic. But it is the free copper ion that is harmful. Concealed within an organic molecule, as in hæmocyanin, the copper atom is innocuous. And lately copper-protein compounds have been found also in mammals, the group of animals to which we belong. There is one such compound in the liver, another in human blood. These, too, are blue, but they are in such small quantities in the body that the colour is invisible.

Now I want to go on to speak of hæmoglobin in crustaceans. It is found in the blood of water fleas and of many others of the small and more primitive members of the group. But no animal – crustacean or other – is known to have both hæmoglobin and hæmocyanin in its blood. Some

crustaceans have one of these pigments, some the other, some neither, but none have both. It is not impossible, however, that a crustacean or a mollusc with both these pigments in its blood may some day be found. For until quite lately no case was known of any animal with two respiratory pigments in its blood. But now we know that the blood of a beautiful marine animal called *Serpula* contains both hæmoglobin and also a green respiratory protein known as chlorocruorin. The result is a brown blood, but the spectroscope has given the secret away. Oxyhæmoglobin and oxychlorocruorin have each their characteristic absorption spectrum and the blood of *Serpula* shows both.

What is an absorption spectrum? Well, the reason why oxyhæmoglobin appears red is that when ordinary white light falls on it, the pigment reflects or transmits only the red end of the spectrum. The blue part of the spectrum is absorbed and stopped. Together with the red, the yellow and some of the green also reach our eye from hæmoglobin, and the effect is the well-known colour of spilt blood. Now, oxyhæmoglobin not only absorbs the blue end of the spectrum, but also blots out two narrow regions in the yellow part. These missing parts of the spectrum are called absorption bands. When the red colour of oxyhæmoglobin is viewed through a spectroscope, the spectral colours are seen, minus the blue and violet, and also minus two narrow parts of the yellow, the so-called absorption bands. These absorption bands are characteristic of oxyhæmoglobin. Other red liquids or solids do not show them. They are situated in the spectrum at very definite wave lengths. This is a fortunate chance for our researches; it is fortunate because these absorption bands enable us to identify hæmoglobin when very little of it is present, for instance in *Serpula*, or in a single water flea. Oxychlorocruorin has its own absorption bands, and those of both pigments are seen in the blood of *Serpula*.

Daphnia is the generic name of water fleas. Several species of the genus *Daphnia* are found swarming in certain ponds,

where they are sought by aquarium keepers to feed their fish. These aquarists know that *Daphnia* may be red or almost colourless (Plates 1, 2), and some think, quite unjustifiably, that red ones are better fish food than pale. Red *Daphnia* may be so numerous as actually to colour the water of a pond Swammerdam, a Dutch naturalist of the seventeenth century, described this phenomenon, saying that 'the water appeared as if changed into blood, which indeed, terrified me at first'. The famous Swedish naturalist Linnaeus wrote in the eighteenth century of *Daphnia*: 'Habitat ubique in aquis dulcibus, tanta sæpe in copia, ut appareat sanguinea'. It was the great English zoologist Sir Ray Lankester who discovered, in 1871, that the red colour of *Daphnia* is due to hæmoglobin. He found this by using a microspectroscope, that is, a spectroscopic eyepiece on a microscope. With this instrument Lankester identified the pigment of *Daphnia*. The hæmoglobin is dissolved in the blood liquid; it is not within corpuscles.

The vast number of *Daphnia* in a single pond may all be red, or pink, or colourless. Only quite lately has it been discovered that these various tints are not due to different varieties or races of *Daphnia*. One and the same individual water flea can become red or pale in the course of its short life of several weeks or months. Whole populations of *Daphnia* change colour in a few days. They do this by synthesizing blood hæmoglobin, or by losing it. Their blood becomes richer or poorer in hæmoglobin. This is quite remarkable, for no other animal is known to gain and lose a respiratory blood pigment at this rapid rate and to this astonishing extent. One asks at once two questions. What difference does the gain and loss make to the life processes of *Daphnia*? And what factors in the pond, or within the body of the animal, cause the synthesis or loss of hæmoglobin? The second query is one with the widest implication; it touches on the causes of hæmoglobin synthesis in ourselves, and of anæmia.

Experimental work in the laboratory has shown that one

all-important factor controlling the redness of *Daphnia* is the quantity of oxygen from the air that is dissolved in the pond water. Living in water saturated, or nearly so, with air, water fleas are colourless. In water with very little oxygen dissolved in it they are red. If red water fleas get into well aerated water, they soon lose hæmoglobin, and in the course of a week become pale pink, and eventually colourless. Looking at this surprising phenomenon from the point of view of utility to the animal, it is perhaps not astonishing that *Daphnia* makes more hæmoglobin for itself in conditions of oxygen scarcity. One is used to marvellous adaptations in living beings. The extra hæmoglobin, with its great affinity for oxygen, should be able to pick up, like a sponge, what little oxygen there is dissolved in poorly aerated water. In this way it should enable the animal to go on living. But this is just a supposition. Some aquatic animals, such, for example, as certain kinds of midge larvæ and pea mussels, are able to live a long time in lake water almost completely devoid of oxygen. These animals have no respiratory pigment – their blood is colourless – so perhaps *Daphnia*, too, could live without its hæmoglobin.

This question has been answered as follows. Pale water fleas were cultured for a week, well fed on single-celled algæ, in water poor in oxygen, until the animals had become red. Then some of them were put into water deficient in oxygen, while other pale ones from the original stock were put into similar water. It was found, each time that this experiment was made, that the red water fleas lived longer than the pale ones. This proves that the hæmoglobin manufactured by *Daphnia* in response to oxygen deficiency is indeed of use to the animal; the extra blood pigment enables it to survive longer when oxygen is scarce.

Not only *Daphnia* behaves like this. Another kind of small crustacean has been found to react to deficiency of oxygen in the same way. This is the brine shrimp, *Artemia*. The brine shrimp, as its name indicates, lives in brine, even in saturated brine, and is found in certain desert pools and in salt

works. It is about two-thirds of an inch long and swims all its life on its back. It is a primitive crustacean, that is, it retains certain anatomical characters of extinct crustacean ancestors. Moreover, it is a defenceless creature that manages to survive only by inhabiting brine, where fishes and insect enemies cannot live. Experiments (B. M. Gilchrist) have shown that *Artemia*, just like *Daphnia*, synthesizes blood hæmoglobin when oxygen is scarce in its surroundings, and it loses hæmoglobin in aerated brine. And, as in the case of *Daphnia*, the respiratory pigment enables *Artemia* to live in conditions of oxygen scarcity which would kill colourless individuals. There are yet other primitive crustaceans which have recently been found to increase the hæmoglobin content of their blood when dissolved oxygen is scarce. Plate 4 is a photograph of one of these, alive. Its name is Apus and it is a rare inhabitant of temporary pools of water in this country.

To return, however, to *Daphnia;* it has been found that hæmoglobin is present not only in the blood of the water fleas but also inside their eggs. This is very strange, for there is no other known case of hæmoglobin in the eggs of an animal, with a single exception. This is the louse, whose eggs may become coloured pink by hæmoglobin sucked in with the blood of the host. Plate 3 shows *Daphnia* with 4 red eggs. Often the colour of the hæmoglobin in the eggs is concealed by a green pigment, but then the microspectroscope reveals its presence by the characteristic absorption bands.

We have as yet no idea where hæmoglobin is manufactured in the body of *Daphnia*. Indeed the site of synthesis is as yet an unsolved problem for all invertebrate animals that have the blood pigment; we do not know in what tissue any of them makes it. In ourselves the hæmoglobin of blood corpuscles is made in the marrow of certain bones. In the invertebrate animals, for all we know, the respiratory pigment may be synthesized in the blood itself, not in cells. In *Daphnia* we do not as yet know where the blood pigment is formed, but we do know the origin of the hæmoglobin

found in eggs. One would naturally have expected it to be synthesized in the ovaries, where the eggs are produced. But this is not the case. The ovaries can be seen clearly in the transparent animals viewed with the microscope, and by watching them carefully it has been discovered (E. I. B. Dresel) that as the cells in the ovaries gradually grow and turn into eggs, they slowly become pink. The spectroscope shows that this colour is due to hæmoglobin. But, strange to say, the origin of the pigment is not in the cells of the ovaries themselves; the hæmoglobin passes over from the blood into the ovaries.

The eggs of *Daphnia* are laid into a brood pouch on the back of the animal. In Plate 3 they can be seen within the brood pouch. In the pouch their development into young *Daphnia* takes about two days; at the end of that time the young swim out and away. Soon after these young are born in this way, within 6 hours of the event, the mother moults. That is, she casts her old skin and acquires a new and bigger one. Immediately afterwards she lays another clutch of eggs in her brood pouch. It is during the few hours before the moult, in preparation for the laying of eggs, that the cells in the ovaries become pink. And at the same time the blood colour becomes visibly paler. Hæmoglobin has passed from blood into future eggs. This in itself is unexpected; it is also surprising when one remembers that the hæmoglobin molecule is a very large one. One is struck by the fact that such a big molecule can pass through a cell membrane. But other such cases are known, for instance the secretion out of cells into our body fluids of hormones and enzymes, which are proteins, and the penetration of virus proteins into cells.

After the eggs of *Daphnia* have been laid in the brood pouch, the blood of the mother gradually regains its previous full hæmoglobin content. These changes in the concentration of blood pigment, a loss before egg laying and a gain afterwards, have been studied quantitatively by a simple method. This enables a numerical value to be given to the amount of hæmoglobin in the blood of a single *Daphnia*. The same

animals with both 'kidneys' red; no doubt such blockage would soon kill them.*

Water fleas make hæmoglobin when they need it, that is, when oxygen is scarce. This is true of the several species of *Daphnia* that live in ponds. The oxygen content of pond water often becomes very low owing to the respiration of bacteria and other microbes swarming in the water. Then the water fleas turn red. But other species of *Daphnia* live at the surface of lakes. Here the water is always well aerated, and the water fleas of these species are colourless in nature. But these lake-dwelling *Daphnia*, if deprived of abundant oxygen in the laboratory, also become pink with newly-formed hæmoglobin in their blood. Thus they have the capacity of hæmoglobin synthesis when stimulated by lack of oxygen, although they do not profit by this gift in nature.

Brine shrimps, as we have seen, have the same faculty of hæmoglobin production in response to oxygen paucity, and so do other crustaceans, as for instance Apus. Man, too, increases the hæmoglobin in his blood, and the number of red corpuscles, under like circumstances. The atmosphere has a constant proportion of oxygen, but the amount of air, and so of oxygen, diminishes on mountains. As a result of this, mountain dwellers have redder blood. They respond to the stimulus of oxygen lack by synthesizing extra hæmoglobin. The same effect can be produced experimentally in rabbits, by keeping them in a chamber under diminished atmospheric pressure, as if they were living on a mountain. Their blood gets redder. On return to a free life at normal air pressure their blood pales again to its original hæmoglobin content.

The lowered oxygen pressure of mountain air acts on the blood which circulates in the lungs of men or rabbits. But the extra hæmoglobin is formed, not in lungs or bloodstream, but in the marrow of bones, far away from the lungs. What is the chemical link between the two? Does the paucity of oxygen in the blood cause contracting muscles, or another

*Compare also page 108 in *Blood Groups*

tissue, to produce some by-product which in its turn, circulating with the blood, induces bone marrow to make hæmoglobin? Or does oxygen deficiency in blood directly stimulate bone marrow? One would like to think that the second and simpler suggestion was right, for science advances by accepting the simplest explanations which fit the facts. One would like to think more than this. Is it not possible that the normal hæmoglobin production in bone marrow, not only the extra production on mountains, is a result of oxygen scarcity? After all, bone marrow is a strange, out-of-the-way place in the body for such an important manufacture as that of the respiratory blood pigment. Why in bone marrow? Could it perhaps be because here the bloodstream is slow, circulating in a backwater, and therefore its oxygen is exhausted? As an automatic consequence, hæmoglobin would inevitably be formed. This would be a pretty generalization.

But apparently we are on the wrong track. Various recent researches point to this. Firstly, blood taken from the supposed backwater in bone marrow is not particularly deficient in oxygen. Secondly, bone marrow cells, grown outside the animal body by the method of tissue culture, which keeps them alive for a long time, continue to produce red cells and hæmoglobin, and such bone marrow cells do not form more hæmoglobin if they are partially deprived of oxygen. And, thirdly, when rabbits are kept in an atmosphere with an extra amount of oxygen, that is to say in artificial air with more than the normal 21% of oxygen, they do not have pale blood. They do not produce less hæmoglobin than usual. They do not respond like red *Daphnia* put into well aerated water.

So, although human beings synthesize extra hæmoglobin in the rarified atmosphere of mountains, just as water fleas do in ponds, and men lose the extra hæmoglobin on return to sea-level, just as *Daphnia* does in aerated water, yet the normal, everyday production of hæmoglobin in our bone marrow is not the result of a paucity of oxygen there. Indeed, man and *Daphnia* are probably similar in this res-

pect. For even when pale or apparently colourless, the water fleas always have a minimal amount of hæmoglobin in their blood, a basic hæmoglobin level. This is true even of the most hyaline, colourless, *Daphnia* species living in the surface waters of lakes. If a large quantity of these water fleas is caught in a silk net, and then crushed, and the juice squeezed through muslin, the resulting liquid is pale pink with traces of hæmoglobin which the spectroscope reveals. These lacustrine water fleas have formed some hæmoglobin without any scarcity of oxygen in the water, or presumably in their blood.

Yet *Daphnia* and other crustaceans, just like a human mountaineer, do manufacture additional hæmoglobin in response to a lack of oxygen. It is at first sight curious that the absence of something, namely oxygen, should produce a positive effect, the synthesis of hæmoglobin. This may, however, mean that when oxygen is present, something that is essential to the synthesis of hæmoglobin is oxidised and so put out of court. What substance could this be? We do not yet know enough about the chemical steps in hæmoglobin synthesis to give even a speculative answer to this question, although in the near future we may be in a better position, since steps in the synthetic process are now being discovered with the help of radioactive isotopes. But there is another possibility regarding the action of lack of oxygen.

Our own bodies are all the time manufacturing hæmoglobin and all the time destroying it: ten million red blood cells are formed and ten million others destroyed each hour. Yet, unless we climb a mountain or become anæmic, the hæmoglobin content of our blood is remarkably constant. What keeps it steady? A hormone, that is a chemical substance circulating in the blood, has long been suspected. Forty years ago, serum from anæmic rabbits was injected into normal ones, which responded by augmenting their hæmoglobin. Anæmia called forth the hormone. But the hormone had no effect on the anæmic rabbits themselves, for their hæmoglobin factory was out of order. Later on

it was shown that guinea pigs, in a rarified atmosphere as if on a mountain, produce the hormone. A human foetus, in the womb, also lives as if on a mountain, for its oxygen supply is precarious and often insufficient. Now blood cannot be taken from a foetus for experimentation, but it is easy to get foetal blood from a human navel string at birth. Such navel blood, too, has a substance in it that stimulates hæmoglobin synthesis; this was proved by injecting it into rats.

Yet these various experiments did not always succeed. A likely reason for this has come to light quite lately. It has been discovered that oxygen destroys the hormone. Serum from the umbilical cord can be guaranteed to increase hæmoglobin production in rats only if collected under paraffin oil, that is, when protected from the oxygen of the air. If such serum is saturated with oxygen, it has no effect. What does all this suggest? It suggests that the hormone is being produced continuously, but that normally most of it is destroyed by oxygen in the blood. Only when oxygen is deficient does much of the hormone persist and stimulate more hæmoglobin synthesis in the bone marrow. In the foetus, or in *Daphnia* swimming in a foul and oxygen-deficient pond, the hormone would be effective. But in our everyday life only just enough of the hormone would survive oxidation to maintain the normal constant hæmoglobin level of the blood.

But there is a fundamentally different possibility. The effect of oxygen deficiency on *Daphnia* may not be on chemical synthetic processes in the cells, or perhaps in the blood, at the place where hæmoglobin is being made. It may be that the lack of oxygen is acting in the water flea's intestine. We know that when dissolved oxygen is scarce in the water in which the animals swim, then in the liquid within their intestine certain oxidised compounds are reduced: ferric iron is reduced to the ferrous state. We know this by a lucky chance. There happens to be present, dissolved in the intestinal fluid of *Daphnia*, an iron-containing

pigment called a hæmochromogen. The microspectroscope shows the characteristic absorption bands of this pigment. We do not yet know how it comes about that this hæm pigment is present in the intestine, but the important point for the moment is that its absorption bands are visible only when the iron which is in its molecule is in the reduced or ferrous state. When it is oxidised to ferric iron, the bands vanish. Well, when we examine our *Daphnia* under the microspectroscope in water poor in dissolved oxygen, we see the hæmochromogen bands within the intestine. As soon as we supply the animal with aerated water the bands vanish, to reappear at once in water poor in oxygen. This shows that in oxygen-deficient pond water, iron becomes reduced from the ferric to the ferrous state within the gut of the water flea.

It is possible that iron compounds in pond water, swallowed by *Daphnia*, can be absorbed freely through the gut wall into the blood only when the iron is in the reduced, that is, ferrous, state. We know, thanks to the hæmochromogen, that a low oxygen content of the pond water does result in ferric being changed to ferrous iron in the intestine. It may be, then, that when pond water is aerated, *Daphnia* is starved of the necessary iron to make its hæmoglobin. Further work must test these various possibilities, which are fundamental to our conception of hæmoglobin synthesis in any animal.

So far we have spoken of the hæmoglobin of water fleas, of brine shrimps and of our own blood as if it were one single chemical substance. But this is not accurate. The hæmoglobin of each kind of animal is chemically different from that of other kinds. In the case of some hæmoglobins – we must now use the word in the plural – this has been known for a very long time. It was the spectroscope that gave away the secret. The existence of the two characteristic absorption bands in the spectrum of oxyhæmoglobin was discovered in 1862. Soon after this it was noticed that these absorption bands have not got precisely the same wave-lengths in the

blood of man and in that of the ramshorn pond snail. The difference is very small, but it was measurable. Evidently the two hæmoglobins are not precisely alike chemically. Since that time many other cases have been found of differences between the red blood pigments of various animals. They are found to crystallize in different forms, they have different molecular weights, and they differ in oxygen affinity, which means the precise low oxygen pressure at which they attach oxygen to themselves.

Thus it came to be known that the hæmoglobins of various kinds of animal differ slightly from one another; the blood pigment of man, horse, dog, rabbit and pond snail could be distinguished. It was further discovered that in one and the same individual there can be more than one kind of hæmoglobin; that the respiratory pigment of blood is not identical with the hæmoglobin in muscle cells that makes meat red. A spectroscope shows the difference, and the two hæmoglobins have different oxygen affinities. Here the difference has an interesting functional significance. Muscle hæmoglobin in ourselves has a greater oxygen affinity than that of the red blood corpuscles of the blood; this assists the transfer of oxygen from blood to working muscle. And the hæmoglobin in the blood of a foetus is different from that after birth. Here again a useful end is served. The higher oxygen affinity of the blood hæmoglobin of the embryo in the womb, compared with that of the mother's blood, helps the transfer of oxygen from mother to foetus.

Since there are different hæmoglobins in different parts of the body, at different times of life and in different kinds of animals, one wants next to know how closely animals may be related to one and yet prove to possess different hæmoglobins in their blood. This is where water fleas have supplied the answer. *Daphnia* constitutes what zoologists call a *genus*. A genus is a group of nearly-related kinds of animal. Within the genus *Daphnia* there is a number of *species*, slightly different from one another. The species are

named *Daphnia pulex*, *Daphnia magna*, and various others. Now, the microspectroscope has shown that the absorption bands of the hæmoglobin in the blood of these various species of *Daphnia* have slightly different wave-lengths. The deduction from this is that they are chemically different. The spectral variations are much less than that between the blood pigments of man and the ramshorn snail, discovered long ago. But they are real and measurable. *Daphnia* is so far the only genus of animals in which this chemical difference between species has been sought and recorded. Such a difference was to be expected, since the differences between species must in the last resort be chemical. But up to now it is only water fleas that have been studied from this point of view.

An interesting side line to the study of differences between hæmoglobins concerns blood-sucking parasites. We know of some of these to our cost, such as mosquitoes and fleas. These insects feed on the blood of a host animal, but their own blood is colourless. There are, however, just a few blood-sucking parasites whose own blood contains hæmoglobin. Such are leeches. In addition to leeches, the only other parasites that feed on red blood and themselves have red blood are certain small crustaceans known to the zoologist as copepods. They are not very distantly related to water fleas. When young they swim freely in the water and later they become permanently attached to the gills of sea fish. There is no English name for them. Their scientific name is *Lernaeocera*. It used to be thought that the hæmoglobin in the blood of these parasites was simply the host's blood pigment sucked into the parasite's stomach and then passed into its blood. But such is not the case. The spectroscope shows that the absorption bands of host and parasite are measurably quite distinct. This means that hæmoglobins of fish and parasite are chemically different from one another. The parasite indeed sucks in hæmoglobin with its blood meal, but it manufactures its own blood pigment.

Since the red blood pigments of different kinds of animal are spectroscopically, and therefore chemically, different from one another, and since this applies not only to different genera, but also to species within a genus, one naturally wants to go a step further than species. One wants to know whether individuals within a species can be shown to have different hæmoglobins. In the end, the differences between you and me must be mainly chemical differences. You and I are composed largely of proteins. And proteins are by far the most complex chemical compounds in our bodies. The small chemical differences between us must be differences mostly of detail in the make-up of our proteins. Can such differences be detected? Hæmoglobin is a likely protein to give an answer to this question. For hæmoglobin is unusual among proteins in having a coloured part of its molecule, a coloured part with characteristic absorption bands in its spectrum which can be measured. So, having established the spectral differences between species of *Daphnia*, the next step was to look for differences between individuals. Water fleas are too small for this research; an individual water flea does not supply enough blood. But rabbits were available and it was found that indeed there were individual hæmoglobins in rabbits. The spectroscope reveals them. In this way individual proteins have been demonstrated.

At the beginning of this article I said that animals might have red, green or blue bloods. The colours are due to respiratory pigments, carrying oxygen in the blood for the animal's respiration. Only four respiratory pigments are known to exist. All are proteins and three of them contain iron. The red pigment is usually hæmoglobin, found in many animals; the green one is chlorocruorin, chemically near to hæmoglobin but only found in the blood of a few marine worms, such as *Sabella* and *Serpula*. A second and rare red respiratory pigment is hæmerythrin possessed by certain marine animals, as *Sipunculus* and *Lingula;* chemically it is not allied to hæmoglobin. The blue pigment is the copper-

containing hæmocyanin in the blood of some molluscs and crustaceans. One naturally asks the question whether such respiratory substances must necessarily be coloured. Why should there not be colourless ones in the blood of some animals? There is, indeed, no reason why respiratory proteins should not exist in the colourless bloods of numerous worms, molluscs and insects. All that can be said is that no such substances have yet been discovered. It is true that they would not advertise their presence as hæmoglobin, chlorocruorin, hæmerythrin and hæmocyanin do. The only way to seek them is to find out whether a vacuum pump can extract more oxygen from a colourless blood than could be accounted for by simple solution in the watery part of the blood. Such researches have not yet unmasked any colourless oxygen carriers.

REFERENCES

Nature, vol. **156**, 1945, pp. 18, 475.
Nature, vol. **160**, 1947, pp. 431, 825.
Nature, vol. **162**, 1948, p. 736.
Nature, vol. **164**, 1949, p. 59.
Proceedings of The Royal Society, B, vol. **135**, 1948, p. 195.
Proceedings of The Royal Society, B, vol. **136**, 1949, pp. 378, 388.

Blood Groups

DR G. FULTON ROBERTS

IT may seem strange that it is dangerous to introduce the blood of one individual into the circulation of another. It is not so surprising that some of the early efforts to save life by transfusing blood from domestic animals into human beings were unsuccessful (see 'The History of Blood Transfusion', *Science News* 3, page 25). With the transference of blood from one human being to another, the results, in the early days, were very varied. On some occasions the patient appeared to benefit from the transfusion, but at other times alarming symptoms and even death would follow.

It soon became evident that all human blood was not the same and that only under certain circumstances was it possible to transfuse human blood into patients safely. The original observations were made by Dr Landsteiner in 1901. He withdrew blood from a number of his laboratory staff and allowed each sample to clot in a separate tube. Under these circumstances the clot sinks to the bottom of the tube, leaving a clear straw-coloured fluid (the serum) lying above. He withdrew the serum from each sample and preserved it, and then added salt solution to the clots so as to form weak suspensions of the red blood corpuscles (which constitute the greater part of the clot). In this way he was able to mix a small volume of each serum sample with an equal quantity of each red cell suspension in all possible combinations. In a number of the mixtures no detectable change resulted, but in others the red cells were found to have stuck together in firm compact clumps which could not be disintegrated easily. This phenomenon is termed agglutination (see Plate 23). As might have been

expected, in no case did the serum agglutinate the red cells which came from the same blood sample, and only in some cases did agglutination occur with other red cell samples.

By analysing these, and subsequent similar investigations, it appeared that agglutination resulted from the interaction of substances in the serum (called antibodies) and substances in the red cells (called antigens). It seems that there are two distinct antigen substances which may be present, together or singly, or may be absent from a person's red cells. So a man may have the A antigen, or the B antigen, in his red cells; or both, or neither.

Similarly, the antibodies occur in such a way that the antibody to A (called anti-A) is never present in the serum of a person who has the A antigen in his red cells, but is always present in those who have not. A similar reciprocal arrangement exists between B and anti-B. Thus there are four main blood groups which, with their incidence in the British white population, are summarised in Table I:

Table I

Red cells contain	Serum contains	Group called	Approximate incidence
Antigen A	Antibody anti-B	Group A	44 per cent.
,, B	,, anti-A	,, B	8 ,, ,,
,, A & B	,, absent	,, AB	3 ,, ,,
No antigen	,, anti-A & anti-B	,, O	45 ,, ,,

It should be said at this stage that the term antibody usually refers to protective substances produced in the circulation as a result of an invasion of the body by a germ or foreign substance. Antibodies are produced regularly in infectious diseases such as pneumonia, and they can be shown to react in several ways with the infecting organism (in this case, the pneumococcus). Such reactions may take the form of agglutination or they may behave otherwise, but in any case the antibodies will react only with the species of organism responsible for the disease, and with none other.

The blood group antibodies, however, occur naturally in healthy people and are not the result of infectious disease, but their properties are parallel in so many respects with the immune antibodies of disease that the terms antigen and antibody have been applied to the blood group phenomena.

There are, thus, four main blood groups, to one of which each human being must belong. From the teleological point of view, it is very interesting to speculate upon the origin and purpose of these groups. It is not easy to see why Nature deems it ill-advised for human beings to mix their blood and contrives to prevent it in this way. It is very remarkable that every person who does not possess the A antigen in his red cells invariably possesses the anti-A antibody in his serum. Blood transfusion is safe between two people of the same group, but it may be fatal to the recipient of blood from a group different from his own. It appears to be an unaccountable facet of the economy of Nature.

Incompatible Transfusion

It is interesting that in considering the compatibility of blood for transfusion, one is concerned only with the antigen. If blood containing a certain antigen is introduced into an individual who carries an antibody to that antigen, severe symptoms of distress will follow. On the other hand, transfusion of blood containing an antibody into a patient whose cells have the corresponding antigen is harmless.

To take examples: transfusing group AB blood (which has no antibodies) into a group A person is dangerous, because the latter has anti-B in the circulation which attacks the donor's cells. However, it is quite safe to give group O blood (which contains anti-A and anti-B) to a group A patient, for the O cells contain no antigen and the donor's anti-A antibodies do not affect the recipient patient's A cells.

One explanation of this phenomenon is that the donor's antibodies are so diluted in the patient's circulation that

they are no longer powerful enough to affect many of his red cells. When he receives an incompatible antigen, however, he is capable of manufacturing as much antibody as may be necessary to attack and destroy all the foreign cells which he has received, with serious consequences. It is the destruction of incompatible cells in the body which causes the symptoms of a severe transfusion reaction.

Another possible explanation of the harmlessness of transfused antibodies is that the tissue cells, as well as the red cells, all contain small quantities of the A and B substances and so the antibody may be neutralised by combining with tissue cells before it can affect the blood cells.

The rules for transfusion are, therefore, that the blood groups of donor and recipient should be the same, but that in emergency group O blood, having no antigens, may be given to anybody irrespective of group (hence the term 'universal donor' applied to this group); while persons belonging to group AB, having no antibodies, may receive blood of any group ('universal recipient').

The symptoms of the transfusion reaction are due to the inability of the kidneys to excrete the destroyed cells and their products. The patient experiences chills, fever and pain in the loins. Usually the urine is discoloured by haemoglobin, the pigment which is liberated from the cells when they are destroyed. In severe cases the kidneys themselves may be damaged by red cell debris and if the general efficiency of the kidneys is seriously affected the patient's life is in danger. Fortunately transfusion reactions are usually mild and transient.

It is quite easy to determine a person's blood group. Blood taken from a group A person is allowed to clot and the serum is withdrawn; it will contain the anti-B antibody. Similarly from a group B person the serum containing anti-A can be prepared. These sera will keep for several years if stored frozen. One drop of the blood sample of unknown group is mixed on a white tile with each of the sera. If the blood is agglutinated by the anti-A serum only,

then the unknown group is A, for only red cells containing the A substance will be affected by the anti-A antibody. The coarse clumps of agglutinated red cells are very easily seen, and will appear in 1-3 minutes (see Plate 21). If the cells are agglutinated by both sera, the unknown group is AB, and if the cells remain unaltered by either serum, then the cells belong to group O (see Table II)

Table II

	Anti-A serum	Anti-B serum
Group A	+	—
Group B	—	+
Group AB	+	+
Group O	—	—

+ = agglutination.

Inheritance

The blood groups are inherited in accordance with Mendelian law, and once fixed, a person's group remains unaltered and unalterable throughout life. When the method of inheritance of these groups is considered, the system appears more complicated than the simple four types so far described. Every inherited characteristic is derived in two parts, one from the mother and the other from the father, and often one part is dominant over the other: so with the blood groups. The child who inherited group A from its father and group O from its mother would have the constitution AO in the red cells, which would appear in tests as group A, since A is dominant over O. If both parents handed on the group A, however, the child's red cells would contain AA, which is also recognised as group A. So that all people who belong to group A (that is, whose cells are agglutinated by the anti-A antibody, from the serum of all group B individuals) are really a mixture of those who are AA and those who are really AO.

Where the two components are the same, the cells are said to be homozygous, but where they differ, one often

being dominant over the other, the term heterozygous is employed. The distinction is an academic one and not of practical importance, since both types of cells behave as group A. Nevertheless it is interesting to consider how the distinction could be made if required. One way would be to determine the blood groups of the parents or children and see if the doubtful group could be deduced; Figure 26 is a hypothetical and extremely unlikely family tree which illustrates the type of deductions which could be made.

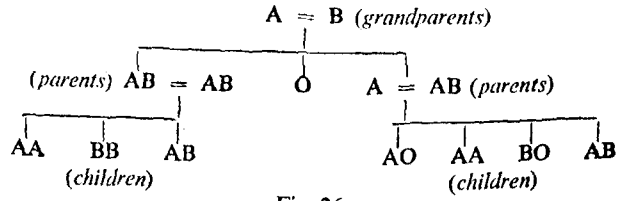


Fig. 26.

The grandparents could be AA or AO and BB or BO respectively. Since one of their offspring is O they must have been AO and BO, for the other combinations (AA×BO, AA×BB, AO×BB) could not give rise to O. The group A parent must also be AO and not AA. If this person subsequently married a person of group AB, the group A offspring could be homozygous or heterozygous, but the group B offspring could be only heterozygous, as shown in the possible children on the right.

If the AB parent, however, were to marry an AB person then the group A and B offspring must be homozygous (the children on the left of the diagram). *It is usually easier to prove a heterozygous group than a homozygous one*; any person of group A or B who has a child of group O must be heterozygous.

Another way, theoretical rather than practical at the present time, would be to test group A cells with an anti-O antibody if this existed. Such an antibody would agglutinate the AO cells (by virtue of the O component) but would not, of course, affect a sample of AA cells. These considerations apply equally to BB and BO cells. Group O cells always have the constitution OO, and group AB is self-explanatory. Which of the two components is handed on by a parent to a child depends on chance and follows Mendelian law.

This simple division into four blood groups, while practical and safe, is not the whole story. It has been found that A is divided into subgroups, A_1, A_2 and A_3. Similarly of course, group AB may be A_1B, A_2B or A_3B. Four-fifths of group A individuals belong to A_1, but A_2 is extremely rare. No subgroups of B have been discovered.

Recent work has also shown that group O does not represent merely the absence of antigens but that there is, in fact, an O antigen which is usually too weak to form antibodies.

Another interesting discovery is that most people secrete the appropriate blood group substances (antigens) in bodily secretions such as saliva and tears. About one-seventh of the population, however, do not do this, and are called 'non-secretors'. In some disorders, the natural outlet of certain bodily secretions may be obstructed so that a cyst may form. Some cysts, in 'secretors', contain the A and B substances in very high concentration, and it has been possible to prepare the blood group substances, chemically pure, from cyst fluid of this kind.

Additional Blood Groups

The ABO system and its sub-types is not the only example of individual differences in human blood; additional blood groups have been discovered. If red cells from a group O person are regularly injected into rabbits or other laboratory animals, antibodies against these cells are produced. The rabbit's serum would be expected to contain an antibody against human O cells. Suppose that this were removed (and this can be done as explained on p. 125) and the serum then tested against a large number of cell samples. If all or none of these samples are agglutinated then no valuable conclusions may be drawn, but it may be that some samples are regularly agglutinated while others are not. If the same proportion of samples is agglutinated every time the test is repeated with a new batch of blood specimens, then it can be confidently deduced that the rabbit serum contains an antibody (in addition to the anti-O one which

was removed) which constantly reacts with the blood of a certain proportion of unselected people. Therefore this antibody is recognising an additional blood group which occurs regularly and is independent of the other blood groups.

It was in this way that the MN system of blood groups was discovered. The population in Britain is divided into M (30%), N (21%) and MN (49%). It has recently been found that another antigen, called S, may be associated with these groups. A person may have in the red cells M or MS; N or NS; MSN or MNS, or any of the combinations of these antigens.

Another and simpler group discovered in this way is P, in which only the alternative forms P and p (i.e. absence of P) are generally recognised.

The MNS and P systems differ in two important ways from the ABO groups. First, there are no naturally occurring antibodies in the serum against MNS or P and thus nothing corresponding to anti-A or anti-B. Therefore no danger ensues from transfusing blood of the wrong MNS or P groups. The second difference is that these antigens are not very powerful, so that transfusions in which these groups are incompatible do not provoke the formation of antibodies. As far as transfusion is concerned, therefore, these groups may be ignored.

The Rhesus Blood Groups

Ten years ago another system of blood groups was discovered which was called the Rhesus factor (see *Science News*, 6, p. 98). In this group also there are no naturally occurring antibodies, but unfortunately the antigen is a powerful one and consequently the rhesus groups occupy a position of importance in the causation of disease.

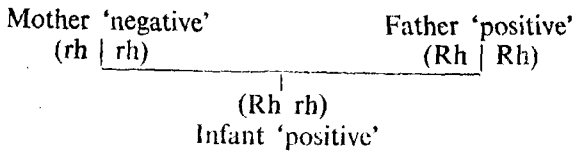
The discovery was made, as with the previous groups, during the course of animal experiments. Blood from the *Macacus rhesus* monkey was injected into rabbits, and the serum from the latter was expected to agglutinate the

monkey's cells. So it did, but also, surprisingly, agglutinated about 85% of a number of samples of human blood. It seemed that 85% of humans possessed the monkey's antigen in their red cells (and were called 'rhesus positive') but the remainder lacked the antigen ('rhesus negative').

The transfusion of rhesus positive blood into a rhesus negative patient is likely to provoke the formation of anti-rhesus antibodies in the recipient's serum, because the rhesus antigen is a powerful one. This, *per se*, may cause no symptoms, and one transfusion of this kind may not be dangerous, but if a second transfusion were given to this patient, the donor's positive cells would be affected by the antibody previously produced, with serious symptoms in consequence.

Here, then, is an important difference between the Rhesus groups (Rh for short) and the ABO groups. In the latter the antibody is present naturally throughout life and so the first incompatible transfusion is dangerous. With the Rhesus groups, on the other hand, the first incompatible transfusion serves merely to provoke the formation of the antibodies (which do not occur naturally) and it is the second or subsequent transfusions which cause symptoms. A patient with anti-Rh antibodies may, of course, be transfused safely with rhesus negative blood, for the antibody has been formed only against Rh positive cells.

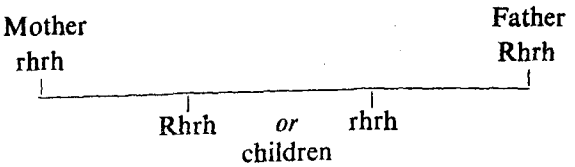
The harmful effects of the Rh blood groups, however, are not confined to transfusion. It must be remembered that during pregnancy the circulation of the infant in the womb is separated from that of the mother only by a very thin membrane, across which various substances needed or discarded by the infant will diffuse. Normally the blood cells of the infant and the mother remain quite distinct and confined to their respective circulations, so that no difficulty arises when the child's group differs from that of its mother. Nevertheless in some cases when a rhesus negative woman gives birth to a rhesus positive infant, which has inherited the Rh antigen from his father



small amounts of the infant's blood may gain access to the maternal circulation and, being incompatible, may provoke the formation of an anti-Rh antibody. This antibody will diffuse back into the infant's circulation and destroy his red cells, with the consequence either that he is born dead or that he exhibits signs of anaemia at birth (haemolytic disease of the newborn), which may be fatal.

It should be emphasised, however, that this does not always, or indeed commonly, happen when rhesus negative women are married to rhesus positive men. In fact, such matings are quite common (about one in ten), but only about one in forty of these marriages is affected. As with transfusion, several stimuli are needed before the disease develops. The first or second infants serve only to provoke the formation of the antibody and are born healthy; the succeeding infants may suffer from the disease.

As explained earlier, the inheritance of the blood groups is of a double nature, so that in this system three basic forms may be recognised; homozygous rhesus positive (expressed thus, RhRh), heterozygous rhesus positive (Rhrh), and rhesus negative (which being recessive must be homozygous, rhrh). This distinction is of some importance in fathers of infants with haemolytic disease, in assessing the probable outcome of any future pregnancies. A rhesus negative woman may be married to a homozygous husband, as shown in the diagram at the top of the page, in which case, once affected, the family must expect all future children to be liable to the disease, for they will all be rhesus positive since the father can hand on only a rhesus positive group. In the case of a heterozygous husband, however, half the children can be expected to be rhesus negative and so escape.



Therefore in an affected family it is important to investigate the father's group in detail so that the future prospects may be assessed.

It is obvious that the inter-relationship of transfusion and pregnancy must be recognised, for the incompatible transfusion of a young female may affect her powers of bearing healthly offspring, and, conversely, those who have given birth to infants suffering from haemolytic disease of the newborn must be carefully treated with compatible blood should they require transfusion at any time. It must be remembered that once anti-Rh antibodies have been produced, they are effective for the remainder of the patient's life.

Lest it be thought from this account that receiving transfusion or bearing babies has now become a hazardous occupation, particularly for rhesus negative people, it should be emphasised that all these troubles are rare. Furthermore, our recent knowledge of these blood groups has almost eliminated any transfusion risks and has materially reduced the mortality of infants from haemolytic disease of the newborn.

Fisher's Genetic Synthesis

So far we have described the rhesus group as divided solely into rhesus positive and rhesus negative. It has been found, however, that a large number of sub-groups exist. Though these have no very great practical importance, it is interesting to consider briefly the story of their discovery, not only to give an idea of the complexity of the blood groups, but also as an example of scientific reasoning.

The first hint of the multiplicity of the Rh groups was

the discovery of a patient's serum which agglutinated all rhesus negative cells and only some samples of rhesus positive ones. It was clear from this finding that rh negativity was not mere absence of the Rhesus antigen but possession of another kind of antigen (for the negative cells must possess an antigen if they are agglutinated by an antibody). Subsequently two more atypical sera were found which agglutinated different proportions of rhesus positive and rhesus negative cells. Presumably these four sera (i.e. the three atypical ones, and the original one made by monkey's cells) were recognising smaller components in the rhesus antigen than had been suspected at first. Eventually eight types of antigenic behaviour were observed among human blood cells which were thought to be sub-groups of the rhesus antigens, and were labelled with symbols such as Rh_1, Rh_2, Rh′, etc.

If a table were drawn up to demonstrate the varying types of behaviour of these cells with the four known rhesus antibodies (or antisera) it would look like this:

Table III

Rhesus antibodies	Rhesus sub-groups Rh_1	Rh_2	Rh_0	Rh′	Rh″	Rh_y	Rh_z	rh
Antibody No. 1.	+	—	—	+	—	+	+	—
*Antibody No. 2.	+	+	+	—	—	—	+	—
Antibody No. 3.	—	+	—	—	+	+	+	—
Antibody No. 4.	—	+	+	—	+	—	—	+

+ = agglutination, — = no agglutination.

*The original Rh antibody which divided humans into Rh+ and Rh—.

It will be seen that each cell type has a special group of reactions with the four antisera, but the most interesting observation from the table was made by Professor Fisher. He noticed that the reactions of the first and the fourth of these antisera were exactly opposite; wherever No. 1 agglutinates a cell type, No. 4 fails to do so, and *vice versa*. He deduced that the antigens recognised by these sera must be mutually exclusive and that a given cell sample must contain one or the other but not both. This led him to

suppose that the two antigens in question were allelomorphs, that is to say alternative inherited characteristics which compete. If a man might inherit blue eyes or black eyes from his parents, he would have one or the other, but not both. So with these two antigens, it was thought that either one (labelled C) or the other (labelled c) would be inherited, but not both.

Antisera Nos. 2 and 3, however, did not have opposite reactions, and so Professor Fisher deduced that they could not be recognising alternative allelomorphs, but that there must, nevertheless, exist two allelomorphs, as yet unrecognised, which would pair up with the two already noticed. He predicted, therefore, that two further antisera would be discovered, in course of time, one of which would have reactions directly opposite to No. 2 (thus recognising antigens which he called D and d) and the other would react opposite to No. 3 (recognising E and e).

This brilliant hypothesis was substantiated by the discovery of the two missing antibodies (in patients having symptoms of incompatible blood transfusion) within a very few years. So the completed table of reactions has a neat and logical appearance:

Table IV

Anti-bodies	Rh sub-groups Rh_1 CDe	Rh_2 cDE	Rh_0 cDe	Rh′ Cde	Rh″ cdE	Rh_y CdE	Rh_z CDE	rh cde
	Respective combination of basic antigens							
Anti-C	+	—	—	+	—	+	+	—
Anti-D	+	+	+	—	—	—	+	—
Anti-E	—	+	—	—	+	+	+	—
Anti-c	—	+	+	—	+	—	—	+
Anti-d	—	—	—	+	+	+	—	+
Anti-e	+	—	+	+	—	—	—	+

It can be seen from the table that the eight different forms of behaviour of the cells with the various anti-Rh sera were really the eight different combinations of the six basic antigens, C, c, D, d, E, e. In their inheritance either C or c is

passed on, but not both, and so with D, d, and E, e. It is not necessary to pursue further the relationship of these complex sub-groups with the original concept of Rhesus positive and negative. The latter concept still remains the most important aspect of the subject, rhesus positivity being synonymous with the possession of the antigen D, and negativity with the possession of its alternative d. Antibodies to the C or E sub-groups are rare and unimportant as causative factors of disease.

Recently it has been found that more than two competitors may exist in some of these sub-groups. In addition to C and c, additional allelomorphs have been found which are labelled c^u, C^w, and C^x. In addition to D and d there are a number (at least twelve) of extra competitors collectively labelled D^u. The Rh system of blood groups, in fact, is of bewildering complexity. The possible competitors for the three sites on the chromosome which carry Rh characters are illustrated below. We shall deal further only with Fisher's original three pairs of allelomorphs, which are given in dark type:

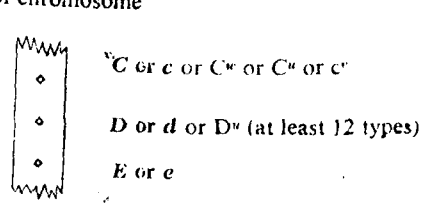


Fig. 27.

Further Genetical Aspects of the Rh Blood Groups

Another interesting though academic aspect of inheritance of the Rh system was described by Professor Fisher. Studying the frequency of the sub-groups in the English population (Table V) he noticed that there were three orders, one of 10 per cent or more, one of about 1-3 per cent, and one very rare.

Normally the antigenic combinations, CDE etc., are inherited as one unit, so that, if the father's group is CDe/CDe and the mother's cde/cde, then the children will all be CDe/cde. Presumably the chromosome, the strip of nuclear material which carries inherited characteristics, has the three sites for the Rh antigens situated very close together.

Table V

CDe	... ...	43.6 per cent.
cde	... ...	37.9 ,, ,,
cDE	... ...	12.8 ,, ,,
cDe	... ...	3.0 ,, ,,
cdE	... ...	1.7 ,, ,,
Cde	... ...	0.8 ,, ,,
CDE	... ...	0.1 ,, ,,
CdE	... ...	0.0 ,, ,,

However, in the early stages of the egg's development the chromosomes sometimes become twisted and even broken, so that the fractured ends may possibly unite with fragments of chromosome with which they were not previously connected. If the C, D and E sites were situated closely together on the chromosome then it is unlikely that a fracture would so occur as to separate them. This doubtless explains why the Rh antigens are inherited as a unit in their original combination.

However, Professor Fisher argued that on rare occasions a break might occur between the Rh sites and that, after rejoining, the combinations might be altered. He assumed that the three common combinations were occasionally affected in this way so that the combinations of the second order of frequency would result.

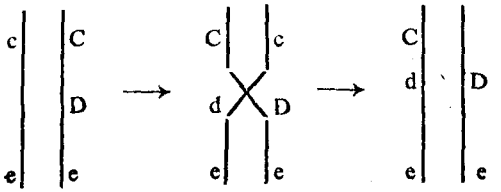


Fig. 28.

In the example illustrated, the cross-over occurs above the level of the D site.

The recombining of the three common groups in this way would result in the four groups of the second order of incidence (Table VI).

Table VI

Crossing over	Result
CDe with cde	Cde and cDe
CDe with cDE	CDE and cDe
cde with cDE	cdE and cDe

It will be noticed that cDe occurs as a result of each combination and would therefore be expected to occur approximately three times as commonly as the other groups of the second order. Reference back to Table V shows this to be the case. CdE, it will be seen, does not result from any of the recombinations but could occur only if one of the groups of the second order were involved. This must be an extremely unusual occurrence and no doubt accounts for the rarity of this group, only one family possessing it having been so far discovered.

The Incomplete Antibody

So far, blood group antibodies have been described as having the property of reacting by agglutination with red cells which contain the specific antigen, and with no others, on the lock-and-key principle. Sometimes, however, an Rh antibody is encountered which 'reacts' specifically with rhesus positive cells but does not agglutinate them. It appears that this form of antibody, called the incomplete antibody, reacts merely by becoming attached to the cells, but does not bind them together. In fact, once rhesus positive cells have reacted with this incomplete antibody they are no longer agglutinable by the ordinary Rh agglutinating antibody. It seems likely that the incomplete antibody occupies the area on the red cell which is concerned with the rhesus antigen, and thus prevents the agglutinating rhesus antibody from

effecting agglutination. A hypothetical and simplified diagram illustrates this point:

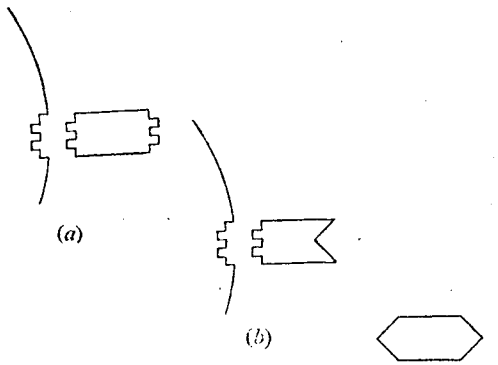


Fig. 29.

(a) On the left is part of the surface of the red cell bearing a configuration which is responsible for the specificity of the rhesus antigen. The agglutinating antibody, shown on the right, fits this configuration at both ends, so that two cells could be drawn together, i.e. agglutinated.
(b) The incomplete antibody on the right will react with the cell, but having done so, blocks the cell against the action of the agglutinating antibody.
(c) A hypothetical shape for Dr Coombs' antibody, which would agglutinate cells which have reacted with the incomplete antibody (see text).

However, it should not be thought that the diagram accurately represents the difference of mode of action of these two types of antibody. In fact, the physico-chemical environment plays an important part and in some circumstances, particularly in the presence of a high concentration of protein, the incomplete antibody can bring about agglutination. The differences are principally a matter of laboratory observation, for in disease the two types of antibody arise in the same way, and produce the same effects; they often co-exist in the same serum.

Dr Coombs has devised an ingenious test to detect

cells which are affected by the incomplete antibody and are unagglutinated. He prepared in rabbits an antibody against the rhesus antibody. The rabbit's antibody, when brought into contact with the affected red cells, will agglutinate them, by virtue of the action of the rabbit's antibody against the rhesus antibody attached to the cells (see Fig. 29).

Treatment

Before leaving the Rh blood groups, a word must be said about the treatment of haemolytic disease of the newborn, the rare consequence of rhesus incompatibility in pregnancy. By means of ante-natal blood tests showing the presence of the rhesus antibodies in the mother's serum, the possibility of the infant being affected can often be prophesied before birth. Sometimes it is possible to cause the infant to be born a few weeks earlier than usual, and thus liberate it from the harmful environment of the maternal antibody. After birth, it may be necessary to tide the infant over the first weeks of life with blood transfusion. The sick child has some of the mother's antibodies in his blood; they will soon be destroyed, but in the meantime they are affecting his Rh positive cells. Until the antibodies are destroyed by natural processes, it may be wise to withdraw some of the infant's blood, and replace it with rhesus negative blood which, of course, is not affected by the antibody. These measures have proved of great value except on the rare occasions when the infant was born dead or moribund; many cases, however, recover without requiring treatment.

Rare Blood Groups

Several further blood groups are rare in the sense that, being such poor antigens, they provoke the formation of antibodies only extremely seldom, and are hardly ever responsible for disease. Consequently only one or two specimens of such antibodies have been found. In some cases the antigen itself is uncommon; only about 8 per cent of individuals, for example, are Lutheran positive, and about 7 per cent are Kell positive.

In two groups (Levay and Graydon) the antigen is so rare that it has been found only among members of the family of the patient in whom it was originally discovered, though several hundred other specimens have been tested.

The Lewis group may prove to be interesting, for not only may it have a number of sub-types, but also it is bound up with the ABO 'secretor' capacity mentioned earlier on p. 111. It seems that all Lewis positive individuals are non-secretors, and all Lewis negative people are secretors, a curious link with the ABO system.

The Uses of Blood Groups

The most important and practical significance of blood groups is for transfusion and pregnancy, and these points have already been fully discussed. It is interesting that the principles of blood group incompatibility in pregnancy have been applied to domestic animals, so that it has been possible to understand more fully a disease called 'icterus' in newborn foals. Here it has been found that the maternal antibody does not gain access to the offspring in the womb, but is transferred in the milk during the first few days after birth. This does not seem to be so important a mode of transmission in humans. These advances in knowledge have resulted in a substantial reduction in the mortality of foals from icterus, an important matter in pedigree stock.

Dr Race whimsically points out that if the claim of Mary Tofts in the early 18th century to have given birth to rabbits had been substantiated, the offspring would have been expected to suffer from icterus due to species antibodies.

Both in research and in other practical fields, an application of the knowledge of blood groups may prove of value. In genetics, for example, the research worker may find blood group inheritance a convenient field for study. The advantages of the blood characteristics are that they are inherited

purely and simply, and are quite uninfluenced by environment. Furthermore, recessive characters, even in the presence of dominant ones, are quite easily recognised, for the heterozygotes can be distinguished from the homozygotes by using the appropriate antibodies. It may be found, for example, that other characteristics are regularly inherited together with a certain blood group, and in this way the groups can be used as 'markers' to follow the genetic behaviour of characters otherwise not easily studied.

Identification by Blood Groups

Landsteiner observed that blood group constitutions would prove as individual as finger prints, and that prophecy will soon be realised. Taking seven of the commoner blood group systems mentioned, it can be calculated that 972,000 possible combinations could occur, of which 23,616 can be recognised with the antisera available at present. Dr Race says that the commonest combination in England is found in only 2 per cent of the population, whereas the rarest combination would be so infrequent that it may never have occurred in the blood of an Englishman. Miss Sanger tested the blood of 250 white people for six of these groups and found that:

133 blood group combinations occurred only once.

29	,,	,,	,,	,,	twice.
10	,,	,,	,,	,,	three times.
6	,,	,,	,,	,,	four times.
1	,,	,,	,,	,,	five times.

She calculated that 53 per cent of those tested were unique in their combination of groups, and 76 per cent were unique or would be confused with only one other from the 250. It can be seen that it may not be long before complete identification of people may be achieved by blood groups.

Even with the limited antisera available at present, identity tests may be used in medico-legal work. In the rare

instance, for example, of unlabelled babies being returned to the wrong cots after feeding, their identity could be established by testing the blood of the mothers and infants concerned, and so the infant could be restored to the correct mother. W. S. Gilbert might have reconsidered the plot of 'The Gondoliers' had he been writing now.

In legal disputes arising over the paternity of infants, the same methods may be employed, in view of the strict inheritance of the blood group characteristics. It is now possible to exonerate a high proportion of men wrongfully accused of paternity.

In certain criminal fields it can be seen that the examination of blood stains will show whether the blood is of animal or human origin and to what blood groups it belongs; it may be possible to assign the origin of the blood unequivocally to the victim or murderer.

The 'secretor' characteristic mentioned earlier makes it possible in some cases to identify the blood group of a person responsible for stains of saliva or (in cases of indecent assault, for example) of semen. Such investigations may play an important part in exonerating or incriminating the suspected person. Of course, in testing stains agglutination is not seen, but the stain will react with the appropriate antisera (and no others) so that they are 'exhausted' of their capacity to agglutinate red cells, when tested subsequently. This process is called 'absorbing out' antibodies, by saturating them with the specific antigen with which they react.

Blood Groups and Ethnology

One of the most interesting aspects of the blood groups is the light they throw on the races of mankind. All the incidences of the various groups which have been mentioned hitherto refer to the population of England, but the relative proportions of the various groups are very different from these in other parts of the world. Group B, for example, is very common in Central Asia, and it may be that the

incidence of this group in Europe is due partly to its introduction by traders and invaders from the East.

It is sometimes possible to draw conclusions from blood group studies of isolated communities which do not intermarry with other local tribes. For example, it is known that Gypsies migrated from India to Hungary about the 15th century; this is supported by their blood group distribution, which accords with that of the pure Indians more nearly than with that of the Hungarians with whom they are settled. By studying the incidence of the A and O groups it can be shown that Iceland, colonised and influenced by the Scandinavians, has in fact to-day blood group distributions very different from those of the Scandinavians. It seems that the Viking invaders were not the most important ancestors of the modern Scandinavians. Iceland, in fact, has a distribution similar to that of Scotland, North Ireland and North Wales, but differing from that of Southern England and Cornwall. It may be possible in time to work out the successive invaders of the British Isles by blood group studies. It seems likely that the early Paleolithic invaders were derived from the same origin as the present Basque civilisation in the Pyrenees. These people have the highest incidence of rhesus negative groups and may represent the original rhesus negative stock. In support of this, Dr Mourant points out that they are the only western Europeans who do not speak an Indo-European language, and that skeletally they have affinities with Crô-Magnon Man. The rhesus negative characteristic is almost unknown in the Far East.

The other groups of invaders are less well demarcated, but it seems that the Iron Age folk and Celts had a high incidence of group O, and the Anglo-Saxons (from North Europe) had a high group A incidence. The Goths from Eastern Europe appear to be responsible for the group B element. Dr Darlington has related blood group distribution and certain speech characteristics in some national groups.

It is only an extension which leads to the testing of mummies from Egypt and America to learn something of blood groups in earlier times. So far the results have been much as historical studies would lead one to expect. It will also prove of interest to determine the relationship of human blood groups to those of the higher primates in an endeavour to fit *Homo sapiens* into his zoological position. This work is only in its infancy, but it appears that the higher apes (chimpanzee, gorilla, orang-utan, and gibbon) have some blood group characteristics in common with man, but that the lower monkeys (except the *Macacus rhesus*) are quite distinct.

It can be seen from this account that a wide field for further investigation remains open. The war-time need for transfusion on a large scale gave an impetus to blood group studies which has resulted in outstanding advances now engaging much attention. It is likely that many more groups will be discovered and that many more interesting results will be obtained.

Those who have mastered the technicalities introduced into this article will have no difficulty in reading further in the following:

'Human Blood Groups': a symposium. *The Advancement of Science*, Vol. 5, No. 20, p. 305, (1949).

'Blood Groups', H. F. Brewer. Chapter IV in *Blood Transfusion*, edited by G. Keynes. (John Wright).

The Rhesus Factor. G. Fulton Roberts. (Heinemann Medical Books).

Some Thoughts on Compound E

H. O. J. COLLIER

PERHAPS it is just a matter of viewpoint, but some scientific discoveries seem to the writer to come as the culmination of a long series of small steps, during which the final aim is clearly in view. The discovery of salvarsan, for example, was of such a type. Ehrlich and Hata patiently tried 606 compounds, gradually getting nearer and nearer the mark until they struck salvarsan. Some scientific discoveries, on the other hand, seem to come as a bolt from the blue. The discovery of the value of Compound E, or Cortisone, in rheumatic conditions seemed to be like this. Yesterday, to all but a few, there seemed no obvious reason to expect that hormones of the adrenal cortex would heal arthritis; to-day the treatment is established. Yet enquiry shows that, after all, this discovery was not entirely unexpected.

Immediately after a discovery is announced an active heartsearching and questioning begins in scientific circles. Why was it made in such and such an institute? What were the conditions essential to its consummation? What were the operative lines of thought? What implications has the discovery for future development?

This questioning has been going on over Compound E, and some interesting things have come to light. For example, at a recent meeting of the Royal Society of Medicine in London it transpired that the hormone by the anterior pituitary gland which stimulates secretion of the adrenal cortex* (that is, the adrenocorticotrophic hormone or ACTH) had been used for treating a case of arthritis in England in

*Further explained on page 132 and 133.

1946. In fact at this meeting the patient, now recovered, who had received it along with other treatment, was asked to trot up on to the platform and demonstrate the ease with which he moved. This ACTH was the very hormone which Hench, Kendall and their colleagues had successfully given to two cases at the Mayo Clinic in 1949. Why had the anti-arthritic activity of ACTH failed to make itself known in this country in 1946 and succeeded in the United States in 1949? The answer given at the meeting was that the dose was too small, the activity of the preparation a little doubtful, and that the supply ran out before full trial could be made. From this we learn, once again, as we found during the war while trying to develop biological knowledge of penicillin in this country, that the factor of supply may well determine the time and place at which discoveries are made. In short, science, like an army, marches on its stomach.

How were the supplies of Compound E and ACTH made sufficient for the crucial experiment at the Mayo Clinic? The supply problem was certainly a teaser. In 1942 Kendall was able to report that he could obtain 340 milligrams of Compound E from 1,000 lb. of adrenal gland. Since about three beef adrenal glands weigh one ounce, we can calculate that 340 milligrams, or 3½ days' treatment for one arthritic patient, demands 24,000 head of cattle if it is to be obtained from adrenal glands.

Clearly the cortisone required for the Mayo Clinic trials would have to come from a source other than the adrenal gland. In 1946 Dr Sarett, in Merck's laboratory, prepared a certain amount of Compound E, but the yield was very, very small. In October of that year, in a conference held in New York, it was decided to try to develop a new method of synthesis, instead of attempting to boost Sarett's original method. By May 1948 the new method had been worked out and was operating, and production has been going ahead ever since.

It is of course well known that this method is not a total synthesis, but a process by which desoxycholic acid, ob-

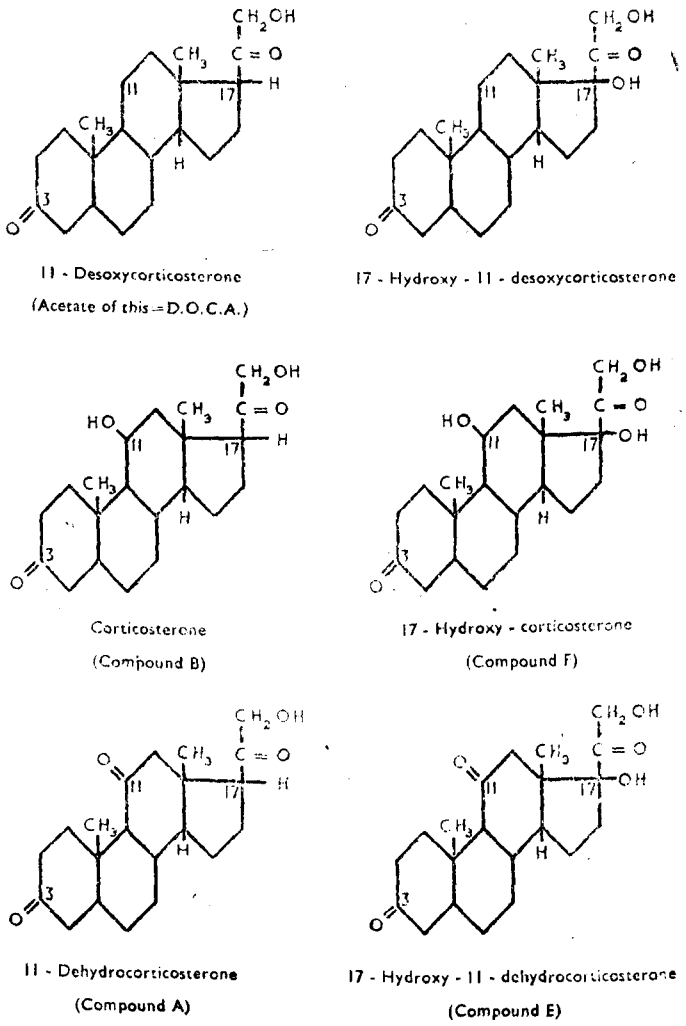


Fig. 30. Chemical structures of six active steroids isolated from the adrenal gland.

tained from ox bile, is converted in about 36 steps to Compound E. While more than 6,000 head of cattle would be required to supply the 100 milligrams of cortisone required by one patient for one day if it were extracted from the adrenal glands, a mere forty cattle are sufficient to supply enough desoxycholic acid for the synthesis of 100 milligrams of Compound E.

It was Kendall who, at the Mayo Clinic in the mid 1930's, had isolated compounds A, B, C, D, E and F (see Fig. 30) from the adrenal cortex, and it was he who naturally played an important part in stimulating and developing the subsequent synthesis of Compounds A and E. The synthesis was worked out at Merck's laboratory in association with Kendall, and consequently we can roughly guess that it was Kendall's isolation of Compound E at the Mayo Clinic in 1935, and his subsequent efforts to arrange its large-scale synthesis, which largely determined that the pioneer work on the treatment of arthritis with this compound would be done at the Mayo Clinic.

There is some possibility of synthesising ACTH, since it is known to be a fairly simple peptide. However, the difficulties of obtaining ACTH from natural sources are somewhat less than those of obtaining Compound E itself. Armours of Chicago estimate that approximately 1,360 hogs are required to produce 10 grams of ACTH. This and other estimates suggest that from 50 to 200 hogs would be required to produce a single daily dose of ACTH. It is easy to see why in the present world position of pig breeding the American continent is likely to be the main producer of ACTH until a satisfactory method of extraction from cattle pituitaries can be employed.

We naturally ask whether the supply of both Compound E and ACTH can be improved. Various plant materials, of which the best known is the seed of *Strophanthus sarmentosus*, a tropical vine, have been suggested as possible sources. It is reported that Sarmentocymarin can be converted into Compound E in about 21 stages. It is also

reported that the yam *Dioscorea mexicana*, which grows in Mexico, New Mexico, Arizona and Texas, produces substances, such as botogenin, which might be used as a starting point for the synthesis of Compound E.

We have seen how the presence of Dr Kendall at the Mayo Clinic was perhaps a determining factor in deciding that the therapeutic possibilities of Compound E should primarily be demonstrated at this Institute. We have seen moreover that until supply difficulties had been overcome the discovery could not have been made, and further that these difficulties were especially likely to be solved at the present time in the United States, where raw materials of animal origin, a powerful chemical industry, and abundant scientific personnel were available. Let us turn to the third question listed at the beginning of this article. What were the operative lines of thought?

Before attempting to describe them, it may be useful to survey briefly the biological facts. The adrenal gland lies, as its name suggests, adjacent to the kidney. It is readily distinguishable into two parts, the rind or cortex, and the core or medulla. The adrenal medulla produces adrenalin, the hormone which brings about changes in the body of a type associated with fear or anger. The cortex, on the other hand, produces the cortical steroids. While the body can dispense with the adrenal medulla and its adrenalin, it cannot dispense with the adrenal cortex and its cortical steroids. Progressive atrophy of the adrenal cortex in man brings on a characteristic condition of weakness which ends in death. The symptoms arising in this process are very definite and characterise the famous Addison's disease. The progress of Addison's disease can be checked by administering suitable extracts of the adrenal glands of animals.

In the daily life of the body the anterior lobe of the pituitary gland, which lies at the base of the brain, regulates the amounts of adrenal cortical hormones which circulate in the blood. The anterior lobe of the pituitary controls the cortex of the adrenal body by means of a special hormone which

it despatches to the adrenal through the blood-stream. This chemical messenger is the ACTH. In turn, the brain influences the pituitary to liberate ACTH. In response to conditions of stress and strain, or to the adrenalin which is liberated by the medulla of the adrenal gland in states of excitement, the brain induces the pituitary gland to liberate more ACTH. And thus in emergencies the body is aided by the protective qualities of the adrenal cortical secretions. This fairly complex circle of relations is perhaps made clearer in Fig. 31.

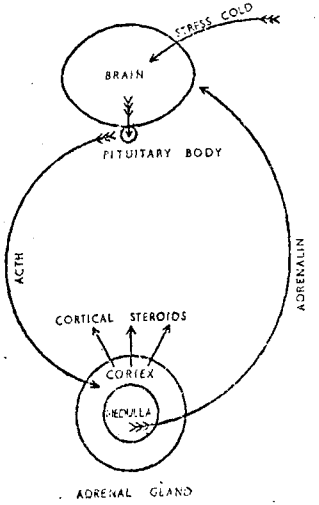


Fig. 31. Diagram of the relations between adrenal gland and pituitary body (*after Thorn and Bayles*).

The fact that Greene had attempted to treat rheumatoid arthritis with ACTH in this country in 1946, while the Mayo Clinic workers were engaged in developing the supply of Compound E for clinical trial in the United States, suggests that medical opinion in general was moving towards the view that the adrenal cortex might be connected with rheumatic conditions.

How did the mental connection between the adrenal cortex and rheumatic conditions arise? The evidence from animal experiments was suggestive. We have already seen that extracts of the cortex protect men and animals against cold and stress. Selye has studied what happens if experimental animals are subjected to continued stress. At first the pituitary reacts and produces ACTH, which in turn stimulates the adrenal cortex. If the stress is maintained longer, the adaptive mechanism becomes overworked and resistance breaks down. Selye was able to produce experimentally in animals various changes similar to rheumatic conditions in man. He concluded that 'the adrenal cortex may play an important role in the pathogenesis of rheumatic and rheumatoid conditions in man'.

It is an interesting point, in the light of Selye's experiments, that in human beings gout and some forms of arthritis are liable to be precipitated or worsened by bodily stress, exposure to cold and so on.

Another line of argument regarding the cause of rheumatoid arthritis has been advanced by Hench. He has pointed out that a patient suffering from arthritis sometimes gets much better during pregnancy or during an attack of jaundice. The beneficial effect of pregnancy in arthritis suggests that the disease might be due to some deficiency in the hormone system. This deficiency might be made good either by secretions of the glands of the fœtus, or by stimulation of some secretion of the mother's ductless glands. This observation certainly seems to narrow the field, so that it can be said that rheumatoid arthritis is related to the hormone balance of the body.

The improvement of arthritis during jaundice is highly interesting. Possibly bile may provide some raw material for the body to make the hormone whose shortage leads to arthritis. If this is so, then it seems that the body and the synthetic chemist might use the same starting point for the manufacture of Compound E. On the other hand, it might be that the jaundiced body excretes steroids less easily than

the normal body, and thus the naturally-produced steroids of the adrenal cortex are conserved.

Experiments and observations such as those mentioned above seemed to point to the idea that the steroids of the adrenal cortex and the cortex-stimulating hormone of the pituitary gland should be tried clinically in intractable cases of arthritis.

Let us now turn to our final question. What are the implications for further development of the discovery that cortisone heals rheumatic conditions? First let us be quite clear that our understanding of the chemistry of the products of cortex of the adrenal gland is far from complete. We have already seen that Kendall isolated the six steroids he called A, B, C, D, E and F from the adrenal cortex. A total of 28 compounds has been isolated from this cortex by biochemists all over the world. Of these, six compounds alone are active as substitutes for the gland when examined by the usual biological tests in animals whose adrenal glands have been removed. These six active compounds are illustrated in Fig. 30. It is interesting to note the pattern of their chemical relationship. Even after the chemist has removed these compounds from the glands there remains an 'amorphous fraction' which possesses strong activity of the same type.

The six active compounds mentioned above differ in the degree to which they affect various organs and systems in the body. What would be the effect of the compounds other than Compound E on rheumatoid arthritis? We know that Compound A is inactive at the doses tried, but we have no information on other compounds. Again, would a balanced mixture of steroids be more active than Compound E alone? As yet we do not know.

We have asked what would be the effect of other cortical steroids in rheumatic conditions. We may also ask what would be the effect of Compound E on other ailments which are associated with stress – psoriasis is perhaps an example.

It is of course well known that treatment of arthritis with cortisone must be maintained or the disease recurs. The question naturally arises as to whether it is possible to give the human body 100 milligrams of Compound E daily for an indefinite period without severe injuries resulting. So far the side-effects have not been severe enough to require the stopping of the treatment. But this question can be answered fully only in the course of time.

It has been estimated that it costs £1,500 to treat a single arthritis patient with Compound E for one week. Is there any material cheaper than compound E which might be used instead? Several lines of enquiry suggest themselves. For example, do any of the substances prepared as stages in the synthesis of Compound E possess any activity against arthritis? Again, will combinations, such as that of vitamin C and DOCA, prove successful?

There is space to consider one final possibility of great interest. It is a fact that administering Compound E causes the number of white cells circulating in the blood to be reduced. It slows down the production of white cells in various glands of the body, such as the lymph glands. Is there then any possibility that Compound E would be beneficial in these cancer-like diseases of the white blood-cells and of the lymph glands – some forms of leukaemia, Hodgkin's disease and lympho-sarcoma?

There is one piece of evidence which bears closely on the above question. In 1944 Heilman and Kendall published a remarkable paper. They described how in 1942 they had successfully transmitted a cancerous tumour from one mouse to another over many generations. They had then dosed some of the tumour-bearing mice with Compound E. The tumours in the dosed animals, instead of growing larger, as they normally did, decreased in size to a remarkable degree. Heilman and Kendall did not publish this work until 1944, because they wished to confirm it and they did not wish to create false hopes. When they did eventually publish, they pointed out that the regression of tumours caused by Compound E did not generally last indefinitely. The tumours usually recurred, and when they did so, they were refractory to treatment with cortisone. This work certainly suggests that the relation of the adrenal cortex to certain forms of cancer might well repay thorough investigation.

REFERENCES

A New Treatment for Rheumatoid Arthritis. Leader, *Brit. Med. J.*, 1949, i, 813.

Greene, R. Royal Society of Medicine, Report of Meeting on Rheumatoid Arthritis. *Lancet*, 1949, ii, 749.

Heilman, F. R. and Kendall, E. C. *Endocrinology*, 1944, **34**, 416.

Hench, P. S., Kendall, E. C., Slocumb, C. H. and Polly, H. F. *Proc. Staff Meet. Mayo Clinic*, April 13, 1949, **24**, 181.

Hench, P. S., Slocumb, C. H., Barnes, A. R., Smith, H. L., Polley, H. F. and Kendall, E. C. *Proc. Staff Meet. Mayo Clinic*, May 25, 1949, **24**, 277.

Kendall, E. C. *Proc. Staff Meet. Mayo Clinic*, May 25, 1949, **24**, 11.

Plant Sources of Cortical Hormones. *Brit. Med. J.*, 1949, ii, 481.

Shoppee, C. W. *J. Endocrinology*, 1947, **5**, 37.

Thorn, G. W., and Bayles, T. B. *The Practitioner*, 1949, **163**, 365.

Some Techniques of Metallurgical Investigation

F. A. FOX

IN the days of our forefathers the term 'metallurgy' if it was used at all meant something very old and simple; it was the art of extracting metals from their ores, and might occasionally be extended to take in their forging and working. To-day the word is almost as general a one as 'engineering', with which it is indeed sometimes combined, in 'metallurgical engineering'. Modern metallurgy borders also on physics and chemistry, mining and mineral dressing, applied mechanics and physical chemistry.

Two very broad divisions may, however, be made, corresponding to the pure and to the applied parts of the subject; these are metallurgical science, and metallurgical engineering. Metallurgical engineering has grown from arts among the oldest known to man, but metallurgical science is still in its infancy, being almost entirely a product of this century. For example, we have been making steel for very many hundreds of years; the provision of a complete scientific account of its properties and behaviour is, however, still far beyond us, and the more it is studied the more complex ordinary mild steel seems to be.

At the beginning of this century the standard tools of the metallurgist included little more than the equipment of the analytical laboratory, the ordinary metallurgical microscope, pyrometers (thermometers), and of course tensile, hardness, and fatigue machines. Since that time, especially with the rise of metallurgical science, many altogether new techniques have been evolved, some of which have already been widely

described, since they are fundamental and general in their application.

Among these are the techniques of X-ray diffraction, which have contributed so much to the understanding of the inner structure of metals and alloys. Microradiography is a much more recent technique whereby pictures, similar to those obtained from an ordinary light microscope applied to an etched metal specimen, may be produced by transmission of X-rays through a thin section of the alloy to be examined. This method is particularly successful where there are large opacity differences between alloy phases, and sometimes yields more information about the alloy than could be obtained by ordinary metallographic methods.

Electron diffraction is a relative new-comer too, but is already well known as a metallurgical tool for the examination of surface films and conditions; neutron diffraction techniques have also recently been announced. In the field of what might be called mechanical metallurgy, electric strain gauges have been of great use in helping to establish with precision what local strains appear when a metal part is stressed.

In what follows I propose to give a few examples of less well-known metallurgical techniques which have been developed in recently published investigations, or which provide new tools of general application.

Metallurgical Engineering

In the production of steel, two very important types of furnace are in use: the open-hearth furnace, and the Bessemer converter. Essentially both of these handle pig-iron, the product of the blast furnace, and convert it by oxidation and other processes into mild steel. In both these furnaces gas reactions are of the utmost importance. In the open-hearth furnace, air/gas mixtures fed in provide the great heat necessary for the maintenance of the furnace temperature, and also provide the appropriate atmosphere for the oxidation of the impurities in the molten bath. In the Bessemer con-

verter, air is blown directly through the molten charge, and direct oxidation of the impurities results.

In order to achieve the best designs and the highest operating efficiency and economy a detailed knowledge of the gas movements is desirable. This has led to interesting recent work on the study of gas flows in these furnaces by means of models.[1] (This and subsequent index numbers refer to journals listed on page 151.) The use of models for the study of aircraft and ships is well known; the application to furnaces is however complicated, because the effects of large changes in temperature and composition of the gases must be taken into account.

In the construction of a transparent plastic model open-hearth steel furnace, it was necessary to choose operating conditions for the model corresponding as nearly as possible to those of full-scale working. Such a requirement does not necessarily depend upon mere geometric similarity. Water was used as the fluid to represent both air and gas, and a 'tracer' technique showed up the flow, using either aluminium powder injected into the water stream, or compressed air, which appeared as a stream of small bubbles along the lines of flow. The flow pattern cannot conveniently be observed by illumination (and photography) of the whole model, since a very confused picture would result from the fact that the flow at different parts may have different direction and velocity. Selected vertical or horizontal planes in the model were therefore illuminated with 'sheets' of light provided by a special optical system.

Similarity of flow in model and in full-scale furnace was achieved by correctly relating the inertia and viscosities; this was done by arranging that the Reynolds number* should be the same for model as for real furnace, both for the fluid travelling along the furnace, and at the 'gas' and 'air' ports. The results show that the gas flow pattern in an open-hearth furnace is very much more complex than had been thought, and that it is likely to be strongly affected by

* See *Glossary*.

quite small changes in furnace design. Furnace wear and tear from hot gas impingement may, therefore, be seriously affected by what could be regarded as an inconsiderable change in the design.

In the work on the Bessemer converter, air was used for producing the blasts in the model, and water or mercury used to represent the steel. Plate 13 shows a typical photograph of the flow pattern in one of the planes of the model side-blown converter, showing particularly the flow of 'gases' above the 'steel' surface.

This model work has shown up such interesting and unexpected effects that work on full-scale furnaces is now going on.

In the fabrication of tubes of copper and brass, the first working operation which is usually put upon the cast piece (the 'billet'), is that of extrusion. In this process a hot cylindrically-shaped casting is put into the container of a hydraulic extrusion press; the central part of the ram of the press (the 'mandrel') pushes downwards and forces a hole

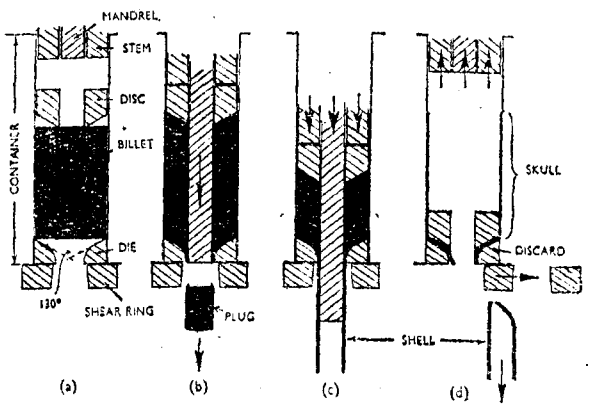


Fig. 32. Successive stages in tube extrusion (*Courtesy of the Institute of Metals*).

right through the billet: the outer parts of the ram then also move downwards and force the rest of the billet out through

the annular space between the die and the still projecting mandrel, thus producing a tube (Fig. 32). The flow of the metal through this annular orifice is obviously somewhat complex; the detail of the flow sequence is, however, important, since for the most economic production of good quality tube it is essential to avoid as far as possible the drawing into the tube of oxide scale from the surface of the billet, an effect which is always liable to take place. The billets are extruded at temperatures of about 675°C. to 1,000°C., depending on the alloy; surface oxidation is, therefore, always to be expected.

In order to follow the flow effects in extruding a billet into a tube, it is obviously necessary to have some form of 'tracer' within the body of the billet, which can be followed through to the tube material, and which can also be observed in the unextruded part of the billet, or 'discard'. In recently published work[2] on this subject the best form of tracer was found to consist of the insertion of metal pins in shallow holes drilled down the side of the billet and in its top face. The pins were made of an alloy resembling the billet alloy (so as not to cause complication of the flow process), but differing in a minor constituent that affected the etching properties of a polished specimen prepared for microscopic examination. For example in experiments with arsenical brass, the pins were made of arsenic-free material, and similarly for deoxidised copper billets the pins were of oxygen-containing copper. The billets were also variegated by making them composite as well as carrying pins; for example, the upper and lower portions were made of one kind of brass, and the middle of another kind.

A technique of this kind permitted the study under practical conditions of the flow of the various copper alloys, and showed up the important fact that the type of flow was governed almost entirely by the friction between the skin of the billet and the press container; other factors were of minor importance. If flow of a lubricated type occurs, then

oxide from the surface is not drawn into the billet to nearly the same extent as happens if non-lubricated flow occurs. It was found that the extrusion problem for the coppers was much easier than for some of the alloys of copper, since the copper-rich oxide layer on them acted as a lubricant between billet and container and led to a lubricated type of flow. This particular type of oxide layer seems to favour slip or slide of the metal under it; the oxide layers formed on copper alloys do not seem to have this property, and the alloys require the aid of external lubricants to give the best results. The pin-insert technique means that from an examination of the structure of the 'skull' left in the container at the end of extrusion, the manner in which the whole billet has flowed during extrusion can be deduced.

Metallurgical Science

It is very well known not only that gases dissolve in metals and alloys when molten, but that they can remain in the solid metal also. The gas in a solid metal may be present merely filling cavities within it; this, however, is not usually so when gas in solid metal is under consideration. What is important, and quite common, is that the gas is truly dissolved in the solid metal or alloy, and consequently no cavities are observable even if the metal is put under the microscope.

The gas is usually regarded as being accommodated on the lattice of the metal by occupying positions in its interstices, rather than replacing metal atoms at some of the lattice points; it is said to be in 'interstitial solution' (compare page 69 in *Aspects of Crystal Chemistry*). Alternatively the gas may enter into combination with the metal or with one of the alloying elements which may be present, forming an oxide, for example, and this compound may then be soluble in the metal.

The point about solid metal containing dissolved gas is that if this metal is heated for any reason, it is very probable that some unpleasant effects will follow from the presence of the gas, which may, for example, tend to come out of solution

and produce cavities at such a pressure as to cause breakdown of the metal at the grain boundaries, or severe blisters.

These general remarks apply to aluminium and its alloys for instance. The gas in aluminium alloys is almost entirely hydrogen, which is often present in amounts markedly in excess of the equilibrium solid-solubility value at room temperature. This hydrogen content of wrought aluminium alloys may be of considerable technical importance if they are to be welded. When a material is welded, there are zones on each side of the line of fusion which are raised to all temperatures intermediate between room and the fusion temperature. In certain aluminium alloys, particularly those containing magnesium as the chief alloying constituent, it is found that if the part being welded has a high dissolved gas content, blisters form near the weld in those regions which have been partly melted. The presence of these blisters is a serious drawback, since they are both weakening and unsightly. Their formation is attributed to the diffusion of dissolved atomic hydrogen and its association at appropriate surfaces or discontinuities within the metal to give molecular hydrogen which may initially, before the blowing up of the blister takes place, be at pressures as high as several hundred atmospheres.

As this effect is a function of the hydrogen content of the solid alloy, the determination of the hydrogen concentration is a matter of technical importance. The hydrogen content is determined by melting in a vacuum, when the hydrogen present in a weighed quantity of the alloy is evolved, drawn off and measured in any convenient manner. This method is accurate and very well established; it requires, however, well-trained operators and rather special apparatus, and is consequently not readily performed in any routine laboratory. There was, therefore, a demand for a simple and quick method of determining the hydrogen content of aluminium alloys with sufficient accuracy for technological control.

It was evident that if specimens of gassy metal were heated to a temperature of semi-fusion, then blisters would

form; further, that if this heating could be done under conditions in which there would be no absorption of hydrogen from outside, and no loss at the surface, then the amount of blistering would be a direct measure of the hydrogen content of the metal. How could such conditions be obtained?

The problem was solved[3] by the discovery that heating the metal in a bath of molten potassium dichromate prevented any loss of hydrogen from the specimen or any gain from outside. This appears to be due to the formation of a heavy brownish-coloured surface film, which seems to be impervious to hydrogen over the experimental temperature range.

After careful checking against vacuum fusion results, it was found that a direct relation existed between the voids formed on a suitable specimen of the test alloy after immersion in molten potassium dichromate at 580°C. for 10 minutes, and its hydrogen content. This relation was established as a calibration curve, from which it was possible to read off for the observed result, which was the ratio of volume of cavities to volume of sound metal, the true hydrogen content. After this work, a relatively difficult vacuum fusion determination became a simple routine control job involving only a determination of the density of the sample before and after heating, the heating being only a matter of immersion in a bath of molten potassium dichromate maintained at 580 ± 5°C. The dichromate adhering to the sample is naturally removed (by boiling water) before the second density determination.

Preferred Orientation

We must turn now from chemical to physical metallurgy. In worked metals, a phenomenon known as 'preferred orientation' is often of considerable technical importance. This effect really derives fundamentally from the fact that the single crystal of a metal is not uniform in its properties in all directions; it is not isotropic. If a single crystal of zinc, for example, is tested mechanically, it is found that there is great ductility if it is pulled in one direction, but brittleness

if it is pulled in a direction at right angles thereto. Even the modulus of elasticity varies very greatly: in zinc the maximum value is more than three times greater than the minimum, the relevant directions being about 70° apart.

In the ordinary pieces of metal of which everyday articles are made the effects of this lack of isotropy of the single crystal vanish, since the enormous numbers of crystals composing them are randomly arranged; no particular crystal directions predominate, and the properties of the piece as a whole are controlled by the *average* properties of the component crystals. However, preferred orientation arises when a piece, though consisting as usual of a large number of crystals, has in some way lost or partly lost its randomness of crystal arrangement. When this happens the piece as a whole begins to revert to the behaviour typical of the single crystal, and to exhibit markedly different properties in different directions.

The principal way in which preferred orientation can appear in ordinary commercial metal products is as a result of inappropriately controlled hot working, and more usually, hot rolling. It leads to unexpected weakness or non-uniformity under stress applied in certain directions. Nevertheless its avoidance is not always easy, and in particular, its recognition by standard methods is not simple. Samples of the material can of course readily be prepared for microscopic examination; however, an ordinarily prepared metallographic sample will give no indication of the grain orientation, which is a function of the inner symmetry of each crystal.

The usual method of detecting preferred orientation is by means of X-rays, when unusually heavy reflexion intensities are obtained from those crystallographic planes which predominate*. This method, well-known though it is, is distinctly slower, more complex and more expensive than that of microscopic examination, which is usually needed anyway as a check on the alloy structure and grain size. This

*See *Aspects of Crystal Chemistry*, page 51.

problem of preferred orientation arises with many wrought materials, among them aluminium and its alloys, and the problem of its rapid detection and determination has been neatly solved[4] for these materials by the evolution of a special technique. Grain orientation might be capable of observation under the microscope if a surface layer of the metal could be transformed into an optically active film – for example by anodic oxidation of the aluminium, and if the bi-refringence of the film on the various grains could be shown up. (Anodic oxidation, or anodising, is a process in which the natural oxide film on the surface of alumium is greatly thickened by making the piece the anode in an oxidising electrolyte; oxygen disengaged at the anode reacts with the metal and builds an oxide film).

For such a method to work it is necessary for several conditions to be met; firstly the film formed must be optically active; there must be a fixed structural relationship between a grain of metal and the film formed on it; the film must be reproducible; it must be possible to determine the optical characteristics of the films and their relationship to the orientation of the underlying metal grains. As a result of research it was found that these conditions could be met, and the technique was perfected in its application to both aluminium and its alloys.

The suspected specimen is first polished and then anodised under controlled conditions in a bath of phosphoric acid with organic and other additions. The anodic film thus formed is found to be uniform over the surface of any one grain or crystal of the metal; however, it changes sharply at the grain boundary and the film on a neighbouring grain has other characteristics. The microscopic observation is carried out in polarised light with crossed nicols, with a magenta tint retardation plate in the system.

The bi-refringence observed on the film varies from grain to grain, and is dependent on the orientation of the film. The effect is that each grain appears with its own colour, which changes through a characteristic colour cycle as the

specimen is rotated. In a specimen which is isotropic – of random orientation – the colours of the grains are also randomly distributed. Where preferred orientation exists, groups of grains of the same orientation are shown up by appearing similarly coloured. Thus the existence of bands or zones of similarly oriented crystals can be detected at a glance under the microscope, without the need for X-ray examination. The method is even capable of extension to give at least semi-quantitative determinations of the orientation of individual grains of the specimen. Aluminium is probably an especially favourable material for a technique of this type, since constancy of relation between the orientation of the metal and the anodic film is a direct outcome of the special volumetric relation existing between the aluminium and its oxide; that is, the oxide occupies about the same space as the metal it replaces. Such a favourable relationship does not always exist; when magnesium is oxidised, for example, an adherent film of magnesia is not formed – it is loose and fluffy.

Hardness

The term 'hardness' is a very familiar one, and the idea of a scale of hardness based on a sequence of standard materials from very soft (talc) to very hard (diamond), each of which will scratch any of the others which come earlier in the list, is also well known. In testing the hardness of metals the method of scratching is not very convenient, and the machines usually employed depend on the measurement of small impressions formed in the surface of the material by an indentor subjected to a known load. Naturally, the harder the material the less deeply will the indentor penetrate, and the smaller will be the measured dimensions of the impression. Of the standard machines in current use, one employs a hardened steel ball as an indentor and the other a pointed diamond; the loads usually range from 1 to 100 kg.

The results are satisfactory for all ordinary routine purposes; however, the hardness figures so obtained (which are

calculated by dividing the load applied by the area of the impression formed), are only average values. Many metallic materials which are of great scientific and technical interest are far from being homogeneous, however, and often contain different constituents which are of a particle size far below that of the indentations made under the loads usually applied. The question may often arise – what is the contribution to the average or gross hardness which is made by each metallographic constituent in a given alloy? Some alloys have hardening or strengthening constituents; how hard are these individually?

It is frequently not possible to measure the hardness of such constituents by obtaining them in massive form and then testing them, since very often they cannot be obtained in the massive state, or even if they could there would be reason to suppose that the material in the massive state would not be truly equivalent to that in the dispersed state in the alloy concerned. The problem really resolves itself into that of producing an extremely fine diamond indentor mounted in such a way as to be highly sensitive under very low loads, which can in some way be brought accurately into contact with selected constituents of the micro-structure of the piece under test. The construction of a micro-hardness testing machine is obviously not an easy matter; a successful answer to the problem consisting of a testing instrument which is to be used in conjunction with a metallographic microscope has, however, recently been described.[5]

One of the critical parts consists of a diamond pyramid-shaped indentor, with the faces worked to an angle of 136° (an angle chosen to agree with that of the indentors of ordinary large-scale instruments), mounted on the front lens of a microscope objective. The specimen can be observed through the objective and the exact part of the micro-structure on which a test is to be made selected, in spite of the mounting of the diamond on the front lens. The mounting is naturally very critical, since the load is transmitted to the specimen through the objective lenses. The load is applied

effectively as dead-loading; the specimen under test is mounted face downwards on a stage up to which the objective lens moves from below. The weight of the specimen itself is counterbalanced by a holder and lever-arm, and the desired indentation load is added directly on top of the specimen. The objective lens is advanced up to the specimen till contact is made and the specimen lifted; the indentation load is then on.

Loads ranging from 5–500 grammes can be applied. Location and measurement of the impression is helped by cross-wires and measuring scales mounted in the eye-piece. With this apparatus, diamond indentations to explore local changes in hardness can be placed on a metal surface with very great accuracy. Plate 14 shows an example: the indentations were placed in the centre of the ruled squares with an error of less than 5μ. Plate 15 shows the varying sized indentations obtained with a constant load (5 grammes) on materials of varying hardness. With special apparatus of this kind closer study will be possible of the parts played by various elements in the micro-structure of many engineering materials.

Thermal Fatigue

As we saw earlier, lack of isotropy in the single crystals of metal can under suitable conditions become, as it were, magnified, so that the whole poly-crystalline piece is affected and material like rolled sheet may have directions of mechanical weakness or non-uniformity which are most undesirable for good service performance. Many other properties besides the purely mechanical are also liable to anisotropy in this way. Among these is the coefficient of thermal expansion, which may be markedly dependent on direction, at least in those metallic crystals which are not cubic in lattice structure. One result of a metal possessing anisotropy of thermal expansion in its single crystal is this: if a piece of the metal which contains many crystals randomly oriented is heated, each crystal will tend to expand chiefly

in a direction different from those of its immediate neighbours. The grain boundary areas will consequently be the sites of conflicting expansions, and local plastic deformation of the metal may occur.

This effect was actually found[6] to occur in some non-cubic metals when they were heated and cooled over even quite small temperature ranges. It was proved to be a consequence of anisotropy of thermal expansion, since the application of similar heat cycles produced no deformation of cubic metals like lead and aluminium (which have no anisotropy of thermal expansion), nor of magnesium which, though non-cubic, shows only a small anisotropic effect.

The extent of the deformation (which is shown up under the microscope by the appearance of striations, the so-called 'slip-lines' – see Plate 16) increases with the number of thermal cycles, and with the temperature range of the cycle. The deformation occurs when the heat cycle involves cooling the piece from room to low temperatures, as well as heating from low to high temperatures. The deformation is absent in single crystals, or in specimens consisting of crystals with a strongly preferred orientation. The grain-boundaries and the general deformation are shown up by careful metallographic work, and slip-lines of the type shown in Plate 16 appear and accumulate in samples submitted to heat cycles.

This general phenomenon is likely to influence the performance of many metal parts; in particular it may effect bearing metals based on the non-cubic metal tin. The behaviour of bearing metals is very complex and difficult of analysis; this hitherto unsuspected effect of 'thermal fatigue', as it has been called, may help to a better understanding of their peculiarities.

(*References overleaf*).

REFERENCES

1. Chesters, J. H., Howes, R. S., Halliday, I. M. D., and Philip, A. R., 1949, *Journal Iron & Steel Inst.*, **162**, 385. Newby, M. P., 1949, *Journal Iron & Steel Inst.*, **162**, 452.
2. Blazey, C., Broad, L., Gummer, W. S., and Thompson, D. B., 1948, *J. Inst. Metals*, **75**, 163.
3. Swain, A. J., 1949 *J. Inst. Metals*, **75**, 163.
4. Hone, A., and Pearsons, E. C., 1948, *Metal Progress*, American Society for Metals, **53**, part 3, 363.
5. Taylor, W., 1948, *J. Inst. Metals*, **74**, 493.
6. Boas, W., and Honeycombe, R. W. K., 1944, *Nature*. **153, 495**; **154, 338**.

Boas, W., and Honeycombe, R. W. K., 1946, *J. Inst. Metals*. **73, 433**.

Natural Science for Non-scientists

IN *Science News 11* and again in *14* we had some discussion on Science teaching which aroused considerable interest. Here, reprinted from *Physics Today*, is an American view of the same subject.

Some sixty teachers of college courses in the physical and biological sciences met at Harvard University on the 9th and 10th of July last year to discuss their aims, methods, and problems in teaching the natural sciences to non-scientists. The five two-hour sessions were devoted respectively to discussions of the methods of science and whether they can be taught, the history of science, methods of instruction, the relationship of natural sciences to the social sciences, and the problem of obtaining suitable teachers.

Possibly because the group was selected for interest in these matters, and possibly because most of the group had met for a similar conference at Princeton a year earlier, there was remarkably little divergence of opinion. At the Princeton conference it had been generally agreed that the new 'block-and-gap' type of course is far more suitable in teaching non-scientists what science is about that the orthodox elementary course in, say, physics, or the general survey course covering a number of sciences (both of which were caricatured unkindly). As its name implies, the block-and-gap course omits whole topics, and lays little emphasis on memorized facts, so that sufficient time may be spent on selected topics to acquaint students with the methods by which scientific understanding is achieved.

From the first session at the Harvard conference it appeared that there is no one 'scientific method' which can

be described briefly and usefully memorized by students. Rather, as Professor Nagel of Columbia put it, examples from the history of science can be used to illustrate the variety of methods which have been used in formulating scientific problems, in formulating concepts to deal with them, in making statements of scientific knowledge, and in evaluating scientific theories or solutions to problems.

Professor Schwab, of the University of Chicago, distributed a paper in which he discusses the diversity of these methods and describes four classifications: taxonomic, such as in the classification of stellar spectra; measurement, such as in atomic or molecular spectroscopy; causal, such as in physiology, and analogical, such as in wave mechanics. Presumably, a well designed course should include examples of as many of these types of methods as possible.

These examples, or 'blocks', or 'case histories' as President Conant calls them, are more successful in teaching students just how science progresses if they can include controversies, or opposing points of view, or theories which have been discarded. Examples mentioned were the Ptolemaic *v.* the Copernican view of the solar system, caloric *v.* the kinetic molecular theory, catastrophic *v.* uniformitarian theories of geology, Newton's particles *v.* Huygens' wave theory of light. There was some feeling that modern developments would make good case histories, particularly in the recently developed sciences of astrophysics, geophysics, and geochemistry.

After this first session there could be little doubt but that the history of science, discussed in the second session, is a valuable adjunct in the teaching of science. It shows the development of scientific understanding, and lends humanistic value to increase the interest of the non-scientists. Unfortunately there are few teachers trained in the history of science, and almost no textbooks. Three recent books of value were mentioned: Sarton's *The Life of Science*, Stimson's *History of the Royal Society*, and F. Sherwood Taylor's *The Alchemists*.

Further discussion of suitable source material was carried on in the session on teaching methods. Selections from original papers by such men as Archimedes, Ptolemy, Copernicus, Newton, Huygens, Leibniz, Dalton, Avogadro, Mendeleyev, Young, Fresnel, J. J. Thomson, Rutherford, Einstein, Shapley and Hubble were reported successfully used. There was no general endorsement of one teaching technique; as Professor Rogers of Princeton put it, the job is to get students to like science by using any or all of the teaching devices available, subject only to the condition that the teaching is sincere.

The posing of broad problems which challenge a student's ingenuity and judgment rather than his mere memorization of fact or a formula for solution, was agreed to be an effective pedagogical device. In fact, some extravagant claims were made for so-called 'discovery experiments', in which the student is encouraged to discover for himself such scientific generalizations as the law of the pendulum, or Ohm's law, or the object-image relationship for a simple lens. It is questionable whether a freshman or sophomore can 'discover' such laws in the same sense that they were first derived; the more cautious attitude at the conference was that such experiments can augment the reading of original papers or other material showing the historical development of the laws and all the difficulties therein.

The kind of broad problem which departs from conventional, 'cook-book' laboratory work, was illustrated by Professor G. E. Owen at Antioch College. His students are asked to consider the advantages and disadvantages of painting a roof with aluminium paint as against some dark colour. In their independent thinking on this problem they can learn a good deal of thermodynamics and radiation theory as well as the experimental methods of testing theoretical conclusions.

A number of social scientists addressed the conference in the fourth session. Skirting the question of whether sociology is properly termed a science, they suggested that introduc-

tory courses in the natural sciences should take pains to eliminate the naïve ideas of some elementary students that natural science can provide answers to all questions. Although many concepts of the natural sciences are not applicable to the social sciences, they felt that training in the natural sciences could develop valuable skills in observation and an appreciation of a scientist's relationship to his data. The latter includes both the process of selecting data for a problem and an awareness of the possible effect of measurement on the data themselves.

Earl J. McGrath, U.S. Commissioner of Education, led off the final session with an appeal for more teachers of science, teachers with better training. There is, of course, a certain difficulty in attracting able men into a low-paid, if stable, profession. More serious, probably, is the schism which is developing between teaching and the modern, specialized, high-pressure research in science. These two activities, teaching and research, which had long been considered necessary concomitants in academic tradition, have recently been separated to such an extent that graduate students now feel they have to make a choice between an academic career of research and one of teaching. In several large physics departments, it was reported, only two or three per cent of the graduate students have any interest in teaching. In contrast to this diminishing supply of teachers, Professor Lark-Horowitz of Purdue reported an increase from ten thousand students in College general-education science courses, to thirty-nine thousand in the United States over the years 1944 to 1947. It is painfully apparent that more interest must be aroused if this teaching of science is to be done adequately.

If there is any popular misconception that the teaching of natural science to non-scientist undergraduates is easy, Professor Philipp Frank of Harvard gave evidence to remove it. In the first place, the general-education course in science generally draws from several fields – usually physics and chemistry, sometimes also astronomy, geology, and

biology. Secondly, the ideal course goes deeply into the methodology, philosophy and history of the scientific problems dealt with. And thirdly, the majority of students in such courses are concentrating on other disciplines. To keep their respect and their interest, the instructor cannot be completely ignorant of his students' fields of specialization. In short, the teachers needed for this effort must have a broad background and a thorough grasp of several fields of science. Graduate students and young Ph.D.s, on whom much of the conventional elementary teaching has been unloaded in the past, are not normally equipped for teaching these courses. However, where a staff of several experienced teachers has organized such a course, some graduate students are reportedly learning as they teach with the group, and gaining a great deal from the experience.

The conference left most of us with the conviction that there is room for considerable improvement in the teaching of science, probably in elementary courses for specialists as well as courses for non-scientists. The growing importance of the natural sciences in our civilization has made it more important than ever before to insure that both the budding scientists and the mass of the population have a balanced idea of what science really is.

A UNIQUE CLUB

Just behind Burlington House (Piccadilly, London) where the Royal Academy and a number of scientific societies, including the Royal Society, have their rooms is a unique scientific club. *The Society for Visiting Scientists* (*S.V.S.*, for short) was created by the British Council in 1944 as a place where scientists visiting London from abroad could get information about British Science and scientists, meet their colleagues socially and informally, and even get a cheap bed for the night. It was hoped that when the war ended other countries would start similar clubs in their capital cities, so that wherever he went a scientist would find an organisation waiting to welcome him and help him – but so far this has not happened, and S.V.S. remains the only place of its kind.

The club has a dining room in the basement where an inexpensive lunch is served on weekdays, and dinner on special evenings when there are meetings. A tiny bar and a large lounge with newspapers and magazines occupy the first floor, while the library and offices are on the second, and the third contains the housekeeper's flat, and additional bedrooms for the use of members. Over 200 scientists, representing 20 different countries, have stayed here in a recent twelve month period.

In addition to providing hospitality, and giving occasional parties for special occasions, the club runs a free information service. The Assistant Secretary keeps an enormous card-index "*Who's Who in Science*" which is kept up-to-date and lists most scientists, their present whereabouts, their interests, their chief publications. With its aid the visitor can be put into touch with the people and laboratories he is most anxious to learn about. About once a month, during the winter, an evening discussion meeting is held, usually on the kind of borderline subject which is avoided by the formal scientific societies. *The Scientific Study of Painting, Science and Precognition, The Scientific Approach to Human Problems in Industry, Science in the U.S.S.R., The Outlook in the History of Science* – these are some of the subjects which have been discussed.

The status of the club is assured by its two patrons, the President of the Royal Society and the Chairman of the British Council; while the Astronomer Royal, Sir Harold Spencer Jones, F.R.S., is its current president, and Professor F. J. M. Stratton, F.R.S., its present secretary, from whom details of the different classes of membership – foreign, British, and junior – can be obtained.

BOOKS RECEIVED

British Intelligence Objectives Sub-committee. Overall Report No. 15: THE FERROUS METAL INDUSTRY IN GERMANY 1939-45 (H.M.S.O. 1949) 4s. 6d.

Ibid. No. 26: ABRASIVES: THEIR MANUFACTURE AND USE IN GERMANY 1939-45 (H.M.S.O. 1949) 6d.

These two specialised technical reports are themselves summaries of a series of other reports of even narrower appeal, and intended as some guide to the wartime work which went on in Germany. It is to be hoped that they will prove useful to scientists and industrialists specialising in these particular manufacturing fields. Glancing through them with our layman's eye we had the impression that compilation and editing had not been ruthless enough. It seems to us that a valuable fact in a report of this sort may only too easily get lost if it is surrounded by trivialities – and surely the precise source of bauxite (page 5) and the failure to synthesise diamonds (p. 6) of report No. 26 are two such unnecessary pieces of information. However, the experts may think otherwise, and these publications are for their good.

THE SOIL AND THE SEA, edited by Trevor I. Williams (Saturn Press 1949) 10/6.

This is a collection of eighteen articles which have previously appeared in *Discovery*, *Endeavour*, *New Biology*, and other periodicals. They deal mostly with applied biology and recent trends under titles such as 'Chemical Stimulants for Crops', 'Seeds and Civilisation', Whales and Whaling', 'Moulds and Tropical Warfare', but include one or two essays which though interesting in themselves appear irrelevant to the scope of this book: 'The Senses of Bats' for example.

The photographic illustrations are a good feature. The typography is unpleasing, the text dazzling to behold. General judgment: a lucky dip miscellany, with some interesting catches.

THE STORY OF ATOMIC ENERGY by Frederick Soddy, F.R.S. (published by New Atlantis, 1949) 20/-.

The author, one of Rutherford's most distinguished collaborators, broadly traces the history of physical science (and tilts at a number of his pet aversions) in an account which is comprehensible to anyone with a slight scientific background. He is most successful when dealing with the great period of atomic development around the beginning of this century. In these chapters he vividly conveys the excitement of scientific thought at that time.

The sections on recent atomic energy developments are, however, disappointing. The manuscript was originally prepared in 1947; since then a great deal of information has been published. By now the informed reader expects more than the Smythe report, but the author makes no attempt to bring his work up to date.

This criticism extends to other fields: the references to cyclotrons and to cosmic rays, for instance. These objections will probably be more important to the physicist than to the lay reader, who, if he does not mind the rather angular quality of the writing, should find here a useful and well-illustrated introduction to modern physics.

P.

THE THEORY OF INBREEDING, by Professor R. A. Fisher (Oliver and Boyd Ltd.), 10/6.

One of the most spectacular successes in the application of science to agriculture has been the use of hybrid maize in the United States where the hybrid stocks have now almost completely replaced the older varieties. The hybrid maize is produced by the crossing together of inbred lines – strains made genetically pure by several generations of self-fertilisation. Many workers are now investigating the possibilities of similar breeding methods in animals. There the lines are made pure by the matings of close relatives. Professor Fisher deals with the rate at which genetic purity or, as it is usually termed, homozygosity, is achieved. The treatment is mathematically rather advanced and the book is quite certainly likely to be very hard going except for those with a mathematical training. The more general question of the possible value of such a breeding method is only briefly discussed, possibly without sufficient reference to the existing experimental evidence on the subject.

Alan Robertson.

GLOSSARY

ALDEHYDE: Any organic chemical compound containing the grouping of atoms – HC=O, which has a number of characteristic chemical reactions enabling it to be identified. (See page 69 for an example).

BIREFRINGENCE: or *double refraction*. Light travels at different velocities through different materials, and when passing from one medium to another, e.g., from air into water, is slowed or speeded, a process technically known as *refraction*. If the medium is not isotropic but has different properties in different directions as happens with some crystals, a light beam passing through it is split up into rays proceeding at different velocities in different directions, and furthermore *polarised* – that is to say with their vibrational energy confined to certain directions in space, instead of being free in all directions as is the case with an ordinary light ray. If these polarised rays of different velocity cross one another's path they may interfere with each other and produce patches of darkness or different colours. Further, polarised light may be accurately recognised by its inability to pass through an apparently transparent substance such as a calcite crystal (usually mounted in a rotatable metal tube and called a *Nicol*) except when the crystal has a certain orientation to the direction of the ray of light. Rotation of the Nicol selects light rays of different polarisation, and so the appearance of a birefringent substance viewed through one changes. A substance which changes the direction of polarisation of a polarised ray passing through it is termed *optically active*.

BLIND SPOT: Nerve fibres are insensitive to light rays, and only specialised cells in the retina of the eye are affected by light. The retina is connected to the brain by the optic nerve fibres, which come into the eyeball in one mass and splay out over the retinal surface to make contact with the receptors. Clearly there must be a 'hole' in the retina to let these optic nerves through, and where there is this hole there are no receptors and consequently a *blind spot*, as you can discover for yourself if you fix your eye on a distant object and move a tiny bit of paper held at arm's length in front of your eye at the same time: there is one position where it suddenly vanishes.

Since light entering the eye has to pass through splayed-out nerve fibres before it can reach the retinal receptors, and these fibres are quite thick, there is one very sensitive region of the retina,

about half a millimetre across, called the *fovea*, where the fibre layer is thinned down to a minimum. When both eyes are used together, they are moved so that the object to be observed is focussed on the fovea of each eye: this movement is called a *fixation*, and the point in space where the axis of one eyeball would meet the other if prolonged is called the *fixation point*.

CHROMATIC ABERRATION: Light rays of different colour or *wavelength* (q.v.) suffer *refraction* (see above) to slightly differing extents in different media. Thus an ordinary glass lens brings the blue rays and the red rays both contained in white light to a slightly different focus, so that a colourless object seen through the lens appears coloured, i.e., suffers chromatic aberration. In optical instruments, the lenses are made of several different materials so that the short-focus of the blue by one is compensated exactly by the long-focus of the other, and the final focus is exact and *achromatic*.

ORBIT (bony): the hollow in the skull, which contains the eyeball.

REYNOLDS NUMBER: is a quotient expressing the relation which the important variables (fluid velocity, density, viscosity) bear to each other in any system of fluid flow.

WAVELENGTH: (λ for short) is a way of describing accurately (quantitatively) the precise colour of a light – the bluer the light the shorter the wavelength, which is measured in μ, thousandths of a millimetre. It is a length because light energy travels discontinuously, in little packets, and the wavelength is actually the distance between one packet and the next. The different wavelengths present in a light beam can be studied by passing the beam through a prism which bends (refracts) the rays of different wavelength to a different extent and so spreads them out on a screen as a *spectrum*, like a rainbow. This apparatus is called a *spectroscope* or if it is equipped to measure the actual wavelengths, a *spectrometer*. (For examples of spectra, see Plate 17; and for some of the uses of spectroscopy, read pages 44, 89 and 100).

About Our Contributors

H. G. J. Collier was born in Natal (Brazil) in 1912. After reading biological and chemical subjects at Cambridge in the early 1930's, he engaged in biological research, with a growing trend towards its medical aspects, at Cambridge, Belfast, Manchester and Liverpool Universities. In 1945 he designed and set up for for Allen and Hanbury a pharmacological research and control laboratory which he now directs. He is the author of numerous papers; of two books, ***An Interpretation of Biology*** (1938), and *Chemotherapy of Infection* (in press); and of a recent television programme. His wife is a radio actress and they have three children.

F. A. Fox graduated in metallurgy at Birmingham and took his first job on a gold mine in West Africa. He has been engaged in metallurgical research ever since he returned to this country and is best known for his publications on magnesium alloys, for which he was awarded his D.Sc. Until recently Deputy Director of the British Welding Research Association, he is now Deputy Technical Manager of the firm of H. J. Enthoven and Sons Ltd., London.

H. Munro Fox, F.R.S., is Professor of Zoology at Bedford College, London University, and editor of *Biological Reviews*. He went to school at Brighton College where, as a day-boy, he could spend his spare time searching for fossils and fishes instead of playing games; he went thence to Gonville and Caius College, Cambridge, where he later became a Fellow. He and his pupils have for a long time studied blood pigments in invertebrate animals. He is the author of *The The Personality of Animals* (a Pelican).

Prof. H. Hartridge, F.R.S. is director of the Medical Research Council's Vision Research Unit and Gresham Professor of Physic at the ancient Gresham College in the City of London. He was formerly Professor of Physiology at the medical College of St Bartholomew's Hospital for twenty years from 1927, and before this he was Senior Demonstrator and Lecturer in the Special Senses at the Cambridge Physiological Laboratory where much of his research was on the functioning of the senses.

D. Hurden was born in 1924, and educated at Uppingham and New College, Oxford, where he read Engineering Science. Since 1944 he has been engaged on combustion research for a well-known firm of aero-engine manufacturers. He is married, but has no children.

William H. Marshall, graduate of Glasgow University, Fellow of the Royal Astronomical Society, Vice-President of the Scottish Branch of the British Astronomical Association, worked on explosives during the war and now teaches mathematics in James Hamilton School, Kilmarnock.

G. Fulton Roberts, who is 29, studied medicine at Cambridge and St. Bartholomew's Hospital. After qualifying, he held several hospital appointments and subsequently worked in the Blood Transfusion Service. He spent two years in the R.A.F. at the Institute of Pathology, and at present is Demonstrator in Pathology in the University of Cambridge.

P. R. Rowland was educated at Bristol Grammar School and Bristol University, where he became a Demonstrator in Chemistry. During the war he did research work for the Ministry of Supply, and he is now a lecturer at Guy's Hospital Medical School, where besides working on crystal chemistry he assists in work on physicochemical problems associated with the oxygenation of the blood.

SUBJECT INDEX

Numbers 1–15

The numbers in brackets indicate the number of the volume.

INDEX OF AUTHORS

Numbers 1-15

Only medical men are named as 'doctor' in this list: the one surgeon is a Fellow of the Royal College of Surgeons (F.R.C.S.). Professors, Fellows of the Royal Society, and holders of civil honours, have been described as such, but no mention is made of any other degrees, doctorates, or academic distinctions. Further information about any author is to be found in the column *About our Contributors*, which appears in each issue.

The numbers in brackets indicate the number of the volume.

About Our Contributors

Crammer is a Cambridge biochemist with a medical training is interested in the simple explanation of scientific discoveries.
. Haslett is a Cambridge graduate, who has made scientific alism his career. He is Editor of *Science Today*, and author number of books of which *Science in Transition* has lately been mmended by the Book Society.
an Ingram Hendey was associated with the British Museum ral History) before the war. In 1940 he went to the Department ival Intelligence at the Admiralty, and spent practically the whole war years in the Far East and Southwest Pacific Area. Is now mber of the Royal Naval Scientific Service and is attached to the al Metallurgical Laboratory, where he is working on problems g to the fouling of under-water structures by marine organisms.
Higinbotham is 26 years of age, was educated at Harrow and ridge, and for a time was a student of Lincoln's Inn. He did ain amount of work in the Ministry of Economic Warfare, and worked with UNRRA. Now farms in Sussex.
Peierls, C.B.E., F.R.S., was born and educated in Germany. it theoretical physics from Sommerfeld, Heisenberg and others. esearch work with Pauli in Switzerland and came to this country 33. After some years spent at Manchester and Cambridge, at Birmingham as Professor of Applied Mathematics (now ematical Physics). Started work with Dr. Frisch on Atomic y in 1940 and did research on the theoretical aspects of this em, first in this country and from 1943 to 1946 in the United . Returned to academic work in 1946, and is now building up a o study fundamental problems.
Perutz. Born in Vienna in 1914. Studied Chemistry at Vienna rsity. Came to Cambridge in 1936 to do research in X-ray llography with Prof. J. D. Bernal. Was appointed research nt to Sir Lawrence Bragg in 1939. Now I.C.I. Research Fellow mbridge University. His main work is concerned with the ular structure of biological systems, particularly the structure crystalline proteins. He pursues glaciology as a side line along climbing and ski mountaineering. Is married and has a three-old daughter.
ele Rabel of Vienna studied Physics and Biology at various rsities including Vienna, Leipzig and Berlin. Author of various cations on the history of science, including a book on *Goethe ant*.
ester Smith, D.Sc. (Lond.), F.R.I.C., obtained his D.Sc. for ch on saponifications of oils and physical chemistry of soap ons. Joined Glaxo Laboratories in 1926, where he has in turn charge of research, development and production of vitamins, etic drugs and penicillin. At present is in charge of biochemical biophysical research.