22

3/61/94

First published November 1951

Subscriptions to
SCIENCE NEWS
and
NEW BIOLOGY
are accepted at
9s each
for four issues
including postage

Made and printed in Great Britain
for Penguin Books Ltd
by The Campfield Press, St. Albans

CONTENTS

PATTERNS OF MUSCLE ACTIVITY IN POSTURE AND MOVEMENT

W. F. FLOYD and P. H. S. SILVER

ELECTRICAL potentials can be recorded from contracting muscles, but a relaxed muscle is electrically silent. This fundamental fact about muscle-action can be made the basis of a method of studying muscle-activity in posture and movement.* The electrical potentials can be led off from the muscles with the aid of suitable electrodes, amplified, and then recorded either photographically or by an ink-writing oscillograph. This technique is called electromyography. Simultaneous recordings made from several muscles show the patterns of co-operative action of the muscles for different movements and postures. The great advantage of the electromyographic method (particularly with surface electrodes on the skin) is that the activity of the muscles can be studied in the intact individual, practically without any interference whatsoever with the normal modes of posture or movement.

The control of muscles is achieved via the nerves which pass from the central nervous system (the brain and spinal cord) to the various muscles of the body. Injury to these nerves results in reduced muscular power: in extreme cases, in complete paralysis. This is what happens in the disease poliomyelitis, otherwise infantile paralysis.

* About 40 per cent of the total body weight in man is muscle which consists of tens of thousands of fine strands or fibres. The individual fibres of a muscle are however too fine to be seen easily with the naked eye. The fibrous appearance of animal meat (which is mostly muscle) is due to large numbers of these muscle fibres adhering together in bundles.

The smallest functional unit of a muscle is a group of muscle fibres all under the control of a single nerve fibre. This unit was termed the motor unit by Sir Charles Sherrington in 1925. When activated by a nerve impulse arriving from the central nervous system, all the muscle fibres of the motor unit contract more or less synchronously, and relax again immediately afterwards unless the contraction is maintained by the arrival of further nerve impulses. The rate of discharge of nerve impulses to the motor unit is one of the factors concerned in the gradation of muscle action. The second factor is the number of motor units which are brought into action for a particular effort. A whole muscle consists of a very large number of motor units, some or all of which can be employed, at different rates of discharge, in order to achieve a desired speed, force and range of movement. Rates of firing of 5 to 50 per second are usual for motor units. More or less motor units in action, together with varied rates of discharge, provide all the variety of action required to grade muscle contraction over the whole range of normal posture and movement.

METHODS OF RECORDING

Each nerve impulse evoking muscle-contraction produces in the muscle an electrical event which is the action potential picked up by the recording electrodes. There are two main ways in which these action potentials can be led off from a muscle and so recorded. In the first method, concentric needle electrodes are inserted into the muscle, and give information about the activity of the part of the muscle in the vicinity of the needle tip. In the second, surface electrodes are placed on the skin overlying the muscle concerned, and give information about the activity of a much greater volume of muscle.

The concentric needle electrode consists of a fine-gauge hypodermic needle (as used for injections) down the bore of which a very fine insulated wire has been inserted, and cemented in position, so that the bare end of the wire just shows at the sharp end of the needle. A second wire is connected to the needle itself. When this specially prepared needle

is inserted (after sterilization) into a muscle it provides metallic contacts with the muscle at two points separated by only a fraction of a millimetre. These contacts or electrodes pick up the electrical activity when the muscle contracts. But, because the two leads are so close, the region of pick-up is localized to a few millimetres around the needle point. Thus, the concentric needle electrode (first described by E. D. Adrian and D. W. Bronk in 1929) provides a means of *sampling* the activity in any part of a muscle. With the needle electrode the activity of individual motor units can be recognized and studied.

Fig. 1 shows the action potentials of several individual motor units as recorded with a concentric needle electrode, inserted into a muscle of the arm in a human subject, and photographed from a cathode ray oscilloscope screen. Note that the potential wave from each of the motor units (distinguished as *A, B, C* in the figure) has its own characteristic size and shape by which it can be recognized throughout the record. The actual wave-shape recorded depends in part on the position of the needle-electrode in the muscle and in part on the size of the motor unit and on its nerve supply. In some of the experiments which will be described we have used needle-electrodes to be sure about the state of activity in the deeper parts of a muscle, as a supplement to the recordings obtained with surface leads.

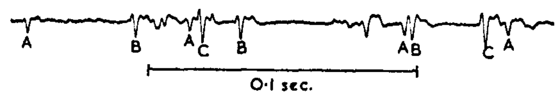


Fig. 1. Action-potentials of individual motor units (labelled *A, B, C*) recorded with a concentric needle electrode from an arm-muscle in a human subject, and photographed from the screen of a cathode-ray oscilloscope

Surface-electrodes consist of any suitable metal contacts which can be placed on the skin. We have used small dome-shaped discs of silver of about 1 cm. diameter. These are electrolytically coated with silver chloride. They are applied to the skin with a cushion of a special electrically-conducting jelly, and held in position with strapping plaster. This gives good electrical contact with the skin and the underlying muscles. With surface electrodes, *all* the electrical events in the muscle contribute to the action potential pick-up, but in varying degree according to their distance from the electrodes, and it is not easy to distinguish the activity of individual motor units. For this very reason, however, the surface leads provide a better index of the activity of the *whole muscle* than do needle electrodes which sample the activity in only one part.

The potentials picked up from muscles in this way amount to a few thousandths of a volt at most, and the smallest recordable activity is a few millionths of a volt. These very small potentials must be amplified before any use can be made of them. This is done with the aid of amplifiers rather similar in design to those in a radio set – the part of it which can be used with a pick-up for playing gramophone records. The amplified version of the muscle activity is applied to an oscillograph which traces out the action potential waves in some visible or recordable form – for example, by means of an ink-writer which traces the waves on moving paper to give a continuous record of muscle action.

TRUNK AND LEG MUSCLES IN STANDING

In the experiments described below we have studied especially the activity of certain trunk and leg muscles, particularly the muscles of the lower part of the back (sacro-spinalis muscles), and the muscles of the side and front of the abdomen (the external oblique muscles in the flank, the rectus abdominis in the midline at the front and the internal oblique muscles in the region of the groin). In the leg the muscles studied were chiefly the calf muscles (soleus and gastrocnemius) and the muscle in front and to the side of the shin bone (the anterior tibial muscle). Fig. 2 shows the approximate positions of these muscles.

In the upright position, when the subject is standing comfortably, there is continuous activity of the internal oblique muscles in the groin, as shown by the continuous action potential discharge (Fig. 3). There may also be seen some activity in

the other abdominal muscles, but it can usually be reduced or abolished by postural adjustments such as swaying; only in the internal oblique muscles is the activity continuously main-

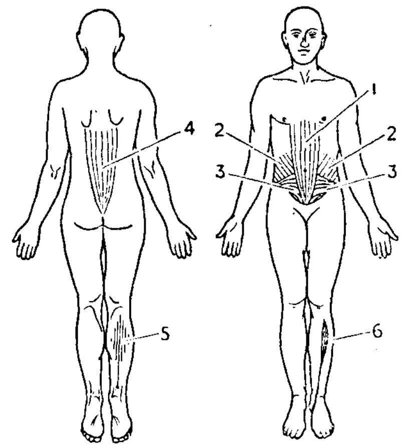


Fig. 2. Sketch to show positions of trunk and leg muscles from which recordings were made. The same number-key is used here and in later figures showing muscle records: (1) recti abdominis; (2) external obliques; (3) internal obliques; (4) sacro-spinalis (also known as erectores spinae); (5) soleus and gastrocnemius; (6) tibialis anterior. The leg muscles (5 and 6) are shown in one leg only in each case, and the two calf muscles, soleus and gastrocnemius, are not distinguished in the figure. In the muscle-records shown in Fig. 4, they are marked as *5a* and *5b* respectively.

tained. These muscles are continually on guard to protect the region of the groin from rupture (hernia). When a straining effort is made, the activity in these muscles increases, thereby providing the greater protection from hernia needed when straining occurs.*

* In a paper in the *Journal of Anatomy*, **84**, 132 (1950), we have shown that this activity can come only from the internal oblique muscles and not from other muscles in the vicinity.

Without any muscle-action at all, the body could, theoretically, be stood upright in a position of balance, just as a pencil can be balanced on end on its point. The position is, however, an unstable one, and the slightest displacement from the position of balance would cause the body to fall over. In practice, there are continuous slight swaying movements in the normal upright position, as will be well known to everyone who has tried to focus on a small distant object with a pair of binoculars. There is difficulty in keeping the object in the centre of the field.

In the normal person, with normal muscle-action to maintain posture, these small displacements are detected by the body through its senses (mainly the eye and muscle sense*). The appropriate muscles are then brought into action, automatically (or reflexly as the physiologist puts it), in order to correct the displacement. These small corrective movements may be brought about by muscles in any part of the body, e.g. the foot, leg, thigh, trunk, arms and neck. In the trunk it is the back muscles and the abdominal muscles which do this. These small postural adjustments are exaggerated when the eyes are closed. As the subject sways forwards and backwards the sacro-spinalis and rectus abdominis muscles fire alternately, the back muscles contracting to arrest the forward movement, and the abdominal muscles contracting to arrest the backward and lateral sway.

In the muscles of the lower leg this intermittent activity to correct slight swaying movements is usually more marked than in the trunk and is more easily seen in the electromyographic records. It is greatly enhanced when the eyes are closed. This is illustrated in Fig. 4, which shows clearly the activity due to swaying movements not only forwards (calf muscles active) and backwards (anterior tibial muscles active) but from side to side with the weight of the body shifting from one foot to the other.

* Muscles are sensitive to stretch and automatically (reflexly) adjust their activity to changes in tension or stretch. E. G. T. Liddell and Sir Charles Sherrington showed this in 1925.

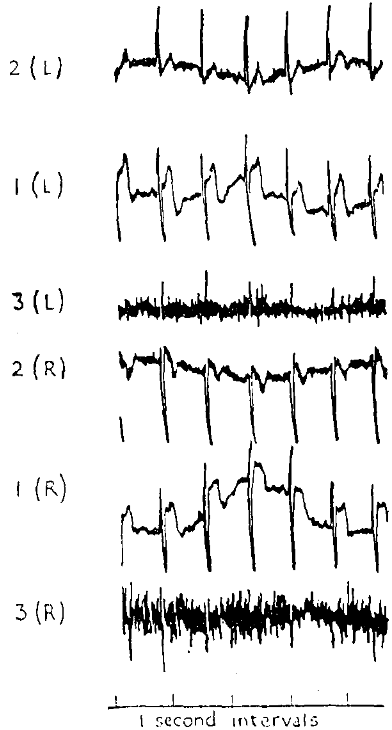


Fig. 3. Electromyogram to show the continuous activity of the internal oblique muscles in the groin in a subject standing in the upright position. Record taken at high amplification (as reproduced, 0.1 inch vertical deflection equals about 23 microvolt), and interrupted about once per second by a large spike due to electrical activity from the heart muscle. In interpreting the record, these ECG (electrocardiogram) spikes are ignored. Muscle action-potentials, originating in the muscles recorded, can be seen clearly between the ECG spikes on the tracings labelled 3(*L*) and 3(*R*), i.e. those from the internal oblique muscles (Fig. 2). There is only a little activity on the other muscle tracings. The letters (*L*) and (*R*) denote left and right muscles. Vertical scale: 0.1 inch = about 23 microvolt.

In the upright stance it is often possible to find a position of balance in which there is minimal, almost zero, activity simultaneously in a great many muscles of the legs and trunk. This position is difficult to maintain for more than a few seconds. It is usually a little easier, however, to maintain a position of the trunk alone (i.e. ignoring, for the moment, the leg muscles) in which both the abdominal and the back muscles show minimal or practically zero activity. This position of balance is one in which the line of the centre of gravity of the trunk (plus arms and head) falls exactly onto the line of the hinge of the pelvis with the legs.

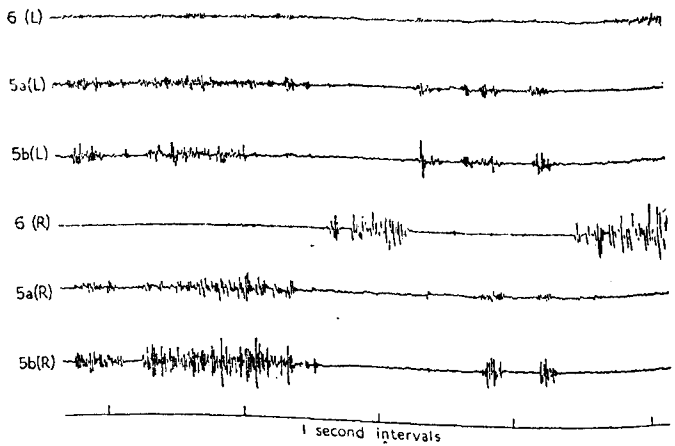


Fig. 4. Electromyogram of leg muscles to show intermittent activity to correct slight swaying movements in postural adjustment. The first part of the record represents a forward sway lasting 2 seconds; only the calf muscles (5*a* and 5*b*) are then active. The subject then sways backwards for just less than 1 second, and only the right tibialis muscle, tracing 6 (*R*), is active (all the weight temporarily on the right foot). The subject then sways forwards again, but more on the left foot than on the right. Finally, the record shows the subject swaying back again, but this time with some weight on the left foot. Note that no ECG spikes are seen. Vertical scale: 0.1 inch = about 90 microvolt.

In slouching, this centre of gravity line falls in front of the line of the pelvis/legs hinge. The body would then topple forwards were it not for the intervention of muscular action to prevent this happening. Fig. 5 shows the activity recorded from the abdominal muscles and the back muscles in a subject standing upright in a balance position (Fig. 5*A*); slouching (Fig. 5*B*); and leaning over backwards (Fig. 5*C*). Notice that to

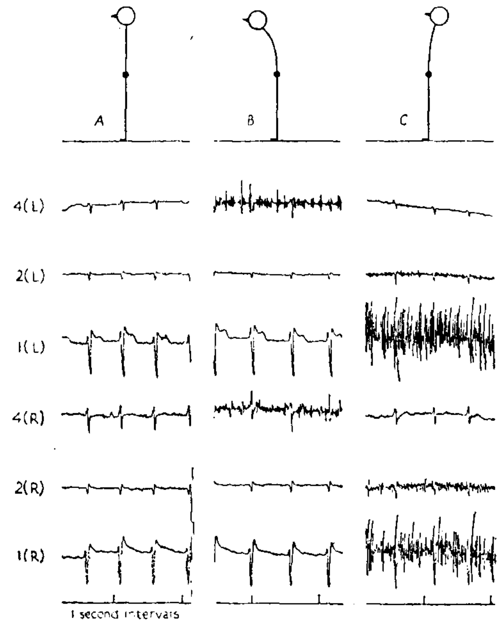


Fig. 5. Electromyograms of trunk muscles to show patterns of activity in the upright position: *A* balanced; *B* slouching; *C* leaning over backwards. In all tracings the ECG spikes are seen, but in tracings taken in the *A* position there is practically no muscle activity. In *B*, only the sacro-spinalis muscles, tracings 4(*L*) and 4(*R*), are active. In *C*, the recti abdominis, tracings 1(*L*) and 1(*R*), are firing vigorously, and there is a little activity in the external obliques, tracings 2(*L*) and 2(*R*). Vertical scale: 0.1 inch = about 180 microvolt.

maintain the slouching posture the back muscles are active. It is these muscles (together with others in the buttock) which contract to prevent the body toppling forwards in this posture. Similarly the muscles at the front of the abdomen (the recti abdominis) become active when the subject leans backwards. The contraction of these (together with certain leg muscles) prevents the body falling over backwards.

These records show that a balance position between these two extremes is an economical one from the point of view of trunk-muscle activity. We may conclude that the upright posture of the soldier (but not the exaggerated sergeant-major attitude of the caricature!) is less fatiguing than the 'scholarly' stoop.

The problem of posture is not quite as simple as this account suggests, however. It is possible to adopt a slouching attitude and then to sway the pelvis forwards and the shoulders backwards and so achieve a posture (rounded shoulders, hollow back and somewhat protruding abdomen) which is still balanced from the point of view of trunk muscle activity, and with little change in the activity of leg muscles. Fig. 6 shows this 'sway back' posture diagrammatically.

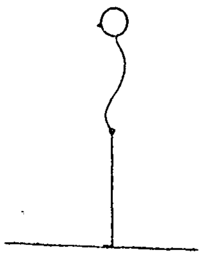


Fig. 6. 'Sway back' posture (see text).

Such a posture has always been condemned as 'bad' but it is certainly not fatiguing as a consequence of muscle activity. Many 'slouchers' actually show considerable back-muscle activity. They do not achieve complete compensation by getting into this 'sway-back' posture.

TRUNK MUSCLES IN BENDING FORWARDS

In bending forwards the centre of gravity of the trunk, head and arms is carried well in front of the line of the pelvis/legs hinge, and the body would topple forwards under the pull of gravity but for the muscle-activity which controls the movement. The muscles mostly concerned are those of the buttocks and back. The movement is usually made as a combination of two major movements: (1) Bending at the hips. This can be done with the back kept straight. It is the movement made in a formal bow. (2) Bending of the trunk only. This can be done without bending at the hips. It is really an exaggerated stoop.

The further forwards the trunk is bent, the greater is the displacement of the centre of gravity of the upper part of the body. Fig. 7 shows the action potentials in the back muscles (sacro-spinalis) during this movement. It is seen that as the trunk comes forwards (Fig. 7, 'flexion starts') the muscle activity increases, corresponding to the increase in the displacement of the centre of gravity. Surprisingly, however, this increasing muscle-activity is not maintained; quite abruptly these muscles relax and become electrically silent (Fig. 7, Record 1, 'critical point') despite the fact that the trunk is not yet bent forwards to its fullest extent (drawing *A*). The man bends further forwards and towards the ground (drawing *B*) without back muscle action (Fig. 7, Record 2, 'full flexion'). These muscles are no longer controlling the movement (the muscles of the buttocks and legs are still in action, however). For the return movement a reversal of the order of events is found. At first the man starts to come up without the help of the back muscles. Then, quite suddenly, the activity recommences (Fig. 7, Record 3, 'critical point') and it continues, with diminishing vigour, until the upright position is reached again.

Mechanics of the back

In order to understand the significance of this relaxation of the back muscles in a posture in which they might have been expected to be vigorously contracting, it is necessary to look at the details of the mechanical relations between the muscles, bones and ligaments of the back. The bones concerned are the back bones or vertebrae. In the small of the back, where these

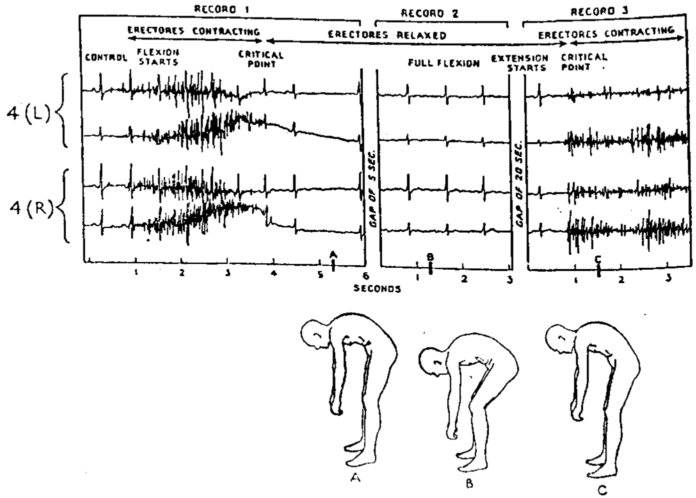


Fig. 7. Electromyographic pattern of activity for sacro-spinalis (otherwise erectores spinae) muscles. Records 1, 2 and 3 are three excerpts from a single tracing obtained while a man bent down to the ground and up again. In the intervals between records, the tracings were identical with record 2. The diagrams of position shown below were traced from photographs taken at the times marked *A*, *B* and *C* at the bottom of the chart. When the man remains bent to the ground (diagram *B*), relaxation continues, as shown in record 2. The contour of the back is the same in diagrams *A* and *B*; only the position of the hips and knees is altered. ECG spikes occur in all three records, and are especially obvious in record 2, in which there is no muscular activity. Record 3 shows that the back muscles suddenly become active again when the trunk is raised. Diagram *C* shows the man's position immediately after the muscles become active again: note the similarity of position between this diagram and that in diagram *A*. Vertical scale: 0.1 inch = about 300 microvolt.

records were taken, the vertebrae have their greatest degree of movement relative to each other. They are mounted one above the other, as shown in Fig. 8, with a cushion of softer tissue (the intervertebral discs) between them. The column is securely anchored at the lower end to the pelvis. The back of the vertebral bones (laminae) are joined together by bands of an elastic-like tissue (ligamenta flava). The vertebral column, made up of these bones, can be bent backwards and forwards, but the amount of bending is obviously limited by these ligaments and the discs. When the trunk is allowed to go forwards from the

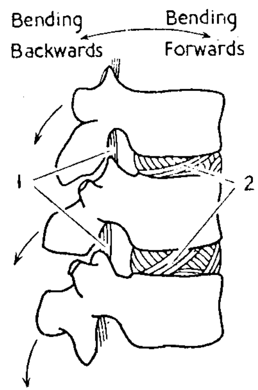


Fig. 8. Diagram of vertebral bones in the small of the back, seen from the right side. Number key: 1. ligamenta flava, the ligaments which join the laminae of the vertebrae together; 2, intervertebral discs, the ligaments which join the bodies of the vertebrae together. The arrows on the left represent the pull of the sacro-spinalis muscles on the vertebral column.

upright position, the weight of the upper part of the body will tend to bend the vertebral column forwards. The mechanics of the system result in the discs being subjected to compression at the front and to tension at the back. The ligaments also become stretched and are in tension.

The sacro-spinalis muscles of the back are in parallel, as it were, with the ligaments. When they contract they exert a pull on the vertebral spines in the downward direction as shown by the arrows in Fig. 8. Contraction of these muscles therefore opposes the forward bending (flexion) of the vertebral column.

Thus when the trunk is allowed to go forwards from the upright position, the weight of the upper part of the trunk starts to flex the vertebral column and to stretch the ligaments and sacro-spinalis muscles. The muscles contract, by reflex adjustments, to sustain the body in the new posture, sharing the load with the ligaments. As the trunk becomes more flexed, more muscle action is required to sustain the posture, and the strain on the ligaments is increased. There comes a position, however, when the ligaments will stretch no further and they take the full strain of the pull of gravity on the trunk. At this stage the muscles relax.

On coming up, the movement is initiated by muscles in the buttocks and legs. At first the back muscles remain relaxed but later contract and take over the strain from the ligaments. Contraction then continues but with diminishing power until the upright position is reached.

TRUNK MUSCLES IN SITTING

In the seated position, when no back support is used, the body is balanced about a fulcrum formed by the line through the two main points of contact of the buttocks with the seat. There are two bony protuberances in this region (the ischial tuberosities) which can readily be felt, and these are the supports for the seated body (cushioned, of course, by the overlying soft tissue of the buttocks). When the legs and feet are fixed on the ground or some rigid support the gravitational pull on the upper part of the body (i.e. all the body above the fulcrum line) is counteracted, as in the upright posture, by suitable contractions of the trunk muscles. In leaning forwards, the back muscles (sacro-spinalis) contract to control the movement. Similarly, in leaning backwards, the abdominal muscles (recti abdominis) contract to maintain the posture. The balanced position, with the numerous small deviations from it, corrected

by appropriate muscle-action, is one of minimal trunk-muscle activity.

Relaxation of the sacro-spinalis muscles can be shown in the seated position also. If the subject 'slumps' in his seat (without back support) the vertebral column becomes fully flexed and the back muscles relax, just as in bending forwards from the upright stance (Fig. 9*a*). Control of the position of the trunk in this slumped-sitting position is achieved by the hip and leg muscles. It is an inelegant posture (Andry, the inventor of the word 'orthopaedics,' said this in 1743) but it is not for that reason a bad posture. Indeed as a position of back-muscle relaxation it is less fatiguing so far as these muscles are concerned. A great many people sit in this 'slumped' position or in a variant of it, in which the back is fully flexed. When a back support is used in a chair, many people slide forward on the seat of the chair and adopt a variant of this position, the back being then fully flexed between the hips and the shoulders. The trunk is leaning backwards but prevented from falling by the back support of the chair. It is the position often favoured by railway travellers (Fig. 9*b*) especially on long journeys and when the seat is wide from back to front: it is more difficult to adopt this posture on a narrow seat.

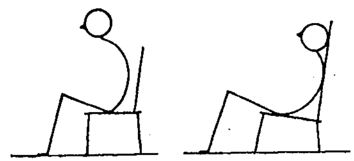


Fig. 9. (*a*) *left*, sitting slumped without back support. (*b*) *right*, sitting slumped with back support. The vertebral column is fully flexed, and the sacro-spinalis muscles are relaxed, in both these postures.

In the slumped sitting posture and its variants, the strain of the weight of the upper part of the body is borne by the ligaments of the back without the assistance of the back muscles. When maintained for long periods the position can become painful, as is well known to car drivers who have to sit for several hours without much change of posture.

WEIGHT LIFTING

When a subject bends forward from the upright position to pick up a weight, even as heavy as 56 lb, the sacro-spinalis muscles of the back are still found to be relaxed in the fully flexed position. Only when the subject begins to straighten the back do the muscles contract. Some subjects do show a small amount of contraction from the moment of taking the strain of the weight, but this is greatly increased only later in the movement. Thus in all subjects, with this method of weight-lifting (Fig. 10*a*), the strain of the weight as well as the strain of the weight of the body (upper part) is borne initially by the ligaments (ligamenta flava and the discs) without or practically without assistance from the back muscles. The tension in the ligaments under these conditions can rise to high values (possibly up to several thousand pounds weight). It is not surprising therefore to find that injuries to the ligaments, especially to the discs, can readily occur during weight-lifting from an incorrect posture, or through incautious movement.

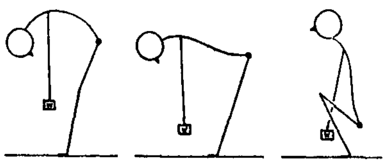


Fig. 10. (*a*) *left*, weight-lifting from upright position, trunk fully flexed and sacro-spinalis muscles relaxed; (*b*) *centre*, weight-lifting from upright position, back straight and sacro-spinalis muscles contracting; (*c*) *right*, weight-lifting with knees bent, back straight, and sacro-spinalis muscles contracting. In (*b*) and (*c*) the strain is shared by the back muscles, and injury to the ligaments is less likely than in (*a*).

There are two ways in which such weight-lifting can be done without imposing so heavy a strain on the ligaments. In the first, the back is kept straighter, and not fully flexed, when the subject bends forwards (Fig. 10*b*) to pick up the weight.

The back-muscles contract vigorously all the time. In the second, the subject bends the legs and picks up the weight between the legs (Fig. 10*c*). In this position it is found that the back is usually straight and the back muscles are contracting vigorously. This is the method used by athletes in weight-lifting competitions. In both of these postures the muscles of the back share the strain with the ligaments and injury to the ligaments is unlikely.

LOAD CARRYING

When a weight is carried in one hand, in the upright posture, it is found that the muscles of the back and some of the abdominal muscles, *on the opposite side*, contract to counteract the effect of the weight. If an equal weight is now put in the other hand, the muscles previously active are found to relax (Fig. 11 *overleaf*). That is, from the point of view of the trunk muscles, it is more economical of muscular effort to sustain a balanced load (equal weight each side) than to carry a load on

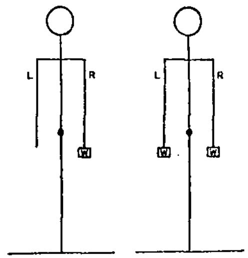


Fig. 12. (*a*) *left*, diagram of back view to show unbalance of one-sided load-carrying; (*b*) *right*, balanced position of load-carrying. In the latter position the centre of gravity of the load and the centre of gravity of the body are practically coincident.

one side only. With the single weight, the one-sided gravitational pull has to be counteracted by trunk muscle action (Fig. 12*a*). With the two weights, their centre of gravity will almost coincide with the centre of gravity of the body, hence very little additional muscular action is required to adjust or maintain the balance of the upright posture (Fig. 12*b*). This is

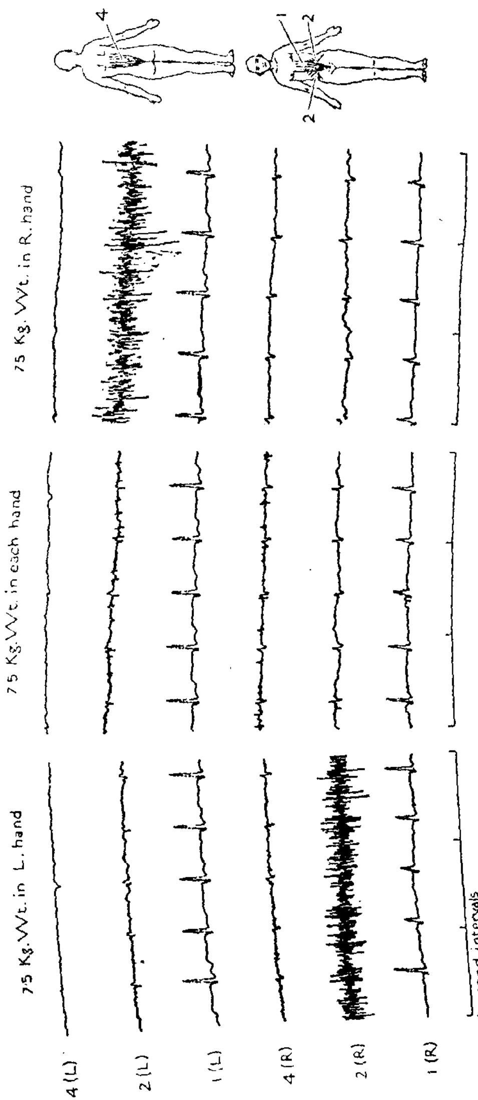


Fig. 11. Electromyograms of trunk muscles to show the effects of equal and unequal distribution of loads carried in the hands: the three records are sections from a single recording. With a 7.5 kg. load in the left hand, the right external oblique muscle contracts; with the same load in the right hand, the left external oblique muscle contracts. A load of 15 kg., equally distributed in the two hands, required less trunk-muscle activity than when 7.5 kg. only was carried on one side. Vertical scale: 0.1 inch=about 160 microvolt.

clearly the principle involved in the Eastern coolie's technique of carrying an exceptionally heavy load balanced on the head. It is also the principle of the yoke, formerly used for carrying milk in pails, one suspended each side, and still in use to-day in the far East for load-carrying.

This result is a matter of common experience, and it is well known that a balanced load is easier to carry than a one-sided load, even when the balanced load is heavier than the single load. The value of the electromyographic method is that it reveals the underlying anatomical and physiological reasons for the difference.

SYMMETRICAL AND ASYMMETRICAL MOVEMENTS OF THE ARMS

Symmetrical movements of the arms, like the symmetrical distribution of the load just discussed, result in minimal activity of the trunk muscles (particularly sacro-spinalis and abdominal muscles). Asymmetrical movements, on the other hand, like one-sided load carrying call for greater trunk-muscle action. If the left arm is lifted sideways the right sacro-spinalis and right external oblique muscles contract – this is strictly analogous to placing a weight in the left hand. If both arms are raised sideways together, these muscles show practically no activity. The function of the trunk muscles in these movements is to fix the trunk and counteract the gravitational forces which act when the limbs are displaced. When the movements are symmetrical, the amount of trunk fixation required is minimal and hence there is minimal trunk muscle action.

REFERENCES

Adrian, E. D. and Bronk, D. W. *J. Physiol.*, **67**, 119 (1929).
Andry, N. *Orthopaedia*, Vol. I, Book 2. London, 1743.
Floyd, W. F. and Silver, P. H. S. *J. Anat.*, **84**, 132 (1950).
Floyd, W. F. and Silver, P. H. S. *Lancet*, **I**, 133 (1951).
Liddell, E. G. T. and Sherrington, C. S. *Proc. Roy. Soc.* '*B.*', **96**, 212 (1925).
Sherrington, C. S. *Proc. Royal Soc.* '*B.*', **97**, 519 (1925).
Weddell, G., Feinstein, B., and Pattle, R. E. *Brain*, **67**, 178 (1944).

PICTURES IN THE MIND

J. A. V. BUTLER

LET us begin by considering how we get our information about the world form our senses: chiefly from sight, hearing, and touch. We know, for example, that when we see an object, the rays of light from the object enter the eye and are focused on the retina, where they cause sensitive cells to send messages or disturbances of an electrical nature along nerve fibres which go to the brain. These disturbances can be traced within the brain to regions which are known to deal with vision, and no doubt the nerve cells which they enter influence many others in the intricate layers of cells which form the cerebral cortex. There our present information comes to an end, although we are aware of the result. Out of the messages we make, and in some way are aware of, a 'picture' of the outside world.

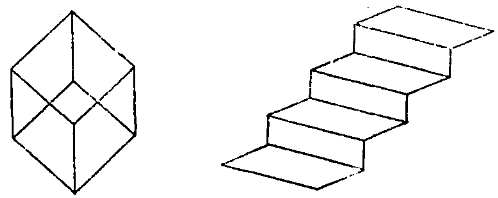
We are so familiar with this picture that we take it to *be* the outside world. We say that we are 'looking' at the world, and primitive and simple people think that when they look out of their eyes and see what is before them, they are performing a positive act. There is no doubt that what we are aware of is a 'construct', made from the varied sensations which reach the brain; a picture which we make by selecting and emphasizing some data and ignoring many others, a picture which in many ways is highly personal.

In certain eye-operations sight has been given to people who have been blind from birth and they have testified to their experience of seeing for the first time. At first they see nothing but confused images – bright lights which they are unable to identify. This may be partly due to their inability to focus on definite objects; but apart from this they cannot recognize the simplest objects from their previous experience. Only slowly, by looking and feeling, do they learn to recognize objects, and

to pick out separate objects in the field of vision by paying special attention to the outlines. They have to learn by trial and error to make a picture of the world out of their sense impressions.

The use which we make of this picture is strictly practical. We use it to guide or plan our actions. When I walk about the room I easily avoid bumping into anything. I am able without any effort to get from one place to another. I sit down in a chair and am not surprised that I can walk about the 'visible' room and sit in a 'visible' chair, but actually I am walking about the 'picture' of the room in my mind and I sit down on the 'picture' of the chair. The picture corresponds so closely with the real world that we accept the one in place of the other, and of course the usefulness of the picture lies in its close correspondence with the outside world.

Usually we have no difficulty in making a picture of our sense impressions, but there are cases in which our sense impressions are ambiguous and may be interpreted in two different ways. If you look at the figures below—



you will see you can make two different 'pictures' from them, which are equally correct interpretations of the visual impressions.

This ability to make a picture of the outside world and to use it in performing actions effectively and accurately is a basic ability of the human brain, on which, as I hope to show, all else depends. It is obviously an ability which is shared by animals, at least to a certain extent. No one who has watched domesticated animals can doubt that they also make some sort of a picture of their surroundings and use it to guide their actions. A horse galloping across a field obviously 'sees' the field and uses his picture of it, as for example when he avoids obstructions and pulls up at the fence. It may be that the human picture is much more accurate and extensive than that of animals.

This method of controlling actions, which the higher animals have evolved, gives them a greatly increased freedom of action, since they can take into account all the elements of their picture of the surroundings. In order to understand what this implies, let us first consider some of the simpler ways in which actions are controlled in living things. Some, like the beating of the heart, go on with very little or no control from the higher centres of the brain. For example, the rhythmic movements of the heart muscles are controlled by nervous impulses which come along the 'motor' nerves in a regular sequence. These impulses are probably started by the heart muscles themselves. It has been found that the muscles possess sensory nerve endings, which send messages to the nervous system to indicate their state of contraction. Thus when the heart-beat reaches a particular stage in its cycle of operations, the nerve endings send out signals to nerve cells in the spinal column, which then initiate the signals to the motor nerves which start the next stage. This method of control, which can be entirely automatic, is commonly used in machines. In engineering language, a part of the output of a process is used to control the input for the next stage. To take a very simple example, we can arrange to have a steady flow of water through a pipe by making its motion control the inflow valve, or we can make a machine perform a repeated sequence of operations by arranging that each step in the sequence shall control the next step, which is what the heart does.

Automatic nervous actions of this kind do not usually involve the higher brain at all. They often continue in animals even when the brain has been removed or disconnected, so that it is inferred that the nerve cells in which the connexion is made between the input and the output are at a lower level in

the spinal cord. But there are often nerves from these cells to and from the higher parts of the brain, which can exert some measure of control when required.

There are also involuntary reflex actions, produced as a response to particular sensations, such as the blinking of the eyelid when a bright light is brought momentarily before the eye, or the sudden withdrawal of the hand when it is touched by a hot object. The bright light starts a signal from the retina to the nerve cells which control the muscles of the eyelid and these are at once contracted quite involuntarily. The sensation of heat stimulates the nerve endings, and these send out signals which produce the sudden contraction of the arm muscles.

It was discovered by Pavlov that reflex actions could also be brought about indirectly. He found that if he rang a bell at the same time as he offered food to dogs, they would quite quickly associate the sound of the bell with food and soon the bell alone would cause them to salivate, etc. This is a case of a sensation of one kind calling forth an action which is not a *direct* response to the stimulus, but is associated with it in a way which has been learnt. Pavlov's followers thought that nearly all actions were conditioned or indirect responses to sensations, although in some cases the connexions between the stimulus and the response might be very indirect and hard to follow.

When we turn to complicated actions we find many features which are ill understood. The mechanism involved in bringing great numbers of muscles into play, each in the right order and to the right degree, is obviously very complicated. The signals which stimulate these muscle movements must reach their destinations at just the right times. The organization of the signals into a coherent pattern is obviously something which is done in the brain; but there are many gradations in the degree of conscious control and in the relation of the action to circumstances.

The ability to perform some actions is clearly innate; the animal is born with it. Chickens can run about very soon after being hatched and will peck at any small bright object. A calf or newly born rabbit is soon able to suck from its mother. Birds can fly, when their wing feathers have developed, with very little or no practice. The patterns of these actions are clearly present at birth; but some stimulus is often required to initiate the action, e.g. no doubt the feel of the mother's breast initiates the sucking (but any soft object will do as well); it has been shown, too, that young birds will gape, i.e. open their mouths for food, whenever a small object approaches the nest. The red-backed shrike has an innate tendency to impale its food on sharp thorns. Actions which are at first innate become subject to modification as experience grows and the usual stimulus (or 'releaser') will no longer call forth an invariable response. K. Z. Lorenz has described his experience with a tame young jackdaw. At first it would come towards him if he did one of three things: he must either move away, more or less quickly, from the bird; or move upwards from where it was; or he must utter a good imitation of the call-note of the species.

> At first the young jackdaw would follow me very willingly ... even when I walked away quite slowly. After a time it became necessary for me to walk away abruptly. After a few more performances I had to add another stimulus, that is I had either to call or to crouch low to the ground beside the sitting bird and jump up instantly. At last, I had to do *all* of these things simultaneously and with the utmost intensity, crouching close to the jackdaw, jumping up suddenly, running away as quickly as I could and yelling jackdaw-calls all the time in order to make the bird take wing at all.

Other types of action clearly have to be learnt. The movements of a kitten are at first disorganized, and it gradually learns how to use its limbs effectively. This seems to be done by trial and error. It will try to repeat actions which produce satisfying results and after many repetitions they will become easy. The nerve connexions required for them are established and the whole sequence can be repeated without error. When such a pattern has been established in the brain, a 'stimulus' will start off the whole series of actions, which is performed as a whole, without any attention being given to the details.

Animals thus acquire a whole repertory of actions which they carry out.

The higher animals are more concerned with the timing and object of an action than with the details of how to perform it. This requires a higher degree of knowledge of the circumstances. The hunting animal for example needs to take a great variety of circumstances into account in order to be successful. No simple scheme of innate behaviour patterns or of conditioned reflex actions will acount for all these actions. A method by which this may be effected was suggested by the Cambridge psychologist, K. J. W. Craig, who died in 1945. His suggestion was that human beings (and possibly animals) use the plan or model of their surroundings which they have made for themselves to plan suitable actions. The brain, he suggested, keeps on trying out possible actions within the plan, and forecasting their results; and when one is found which appears to give a satisfactory result, and which does not conflict with any general rules of conduct the animal has learnt, the action follows. This idea suggests also how the matrix of nervous signals required to bring about the necessary muscular movements is assembled. 'Trying out' an action involves making tentatively all the nervous connexions which will be required if the action is to be effected. Many of the necessary connexions will already exist because they belong to a series of muscular movements which has been learnt. The animal has thus only to select from its 'repertory' a group of movements, of which it can predict the usual result. It makes its selection within the model or plan of the circumstances confronting it.

This method enables the animal not only to make its behaviour suit a wide variety of different circumstances and to carry out effective actions, i.e. actions which achieve their object, in the particular circumstances which confront it. It is able to do this, not only because the action is planned within its picture or model of the surroundings, but also because the signals which bring about the necessary muscular movements are closely connected with this model.

Such a scheme, which allows of forecasting the probable results of actions and also ensures that the actions are closely related to the outside world, gives animals which can use it a freedom and effectiveness of action they could not otherwise have. They can move about in the world of autonomous creatures having a control over their actions and, because they can choose between possible actions and select that which seems best, a certain amount of 'free will'. This is the beginning of individuality and personality and must have been a major step in the elaboration of intelligent sentient creatures.

When we consider the nature of this model or plan, however, we see that it differs in many respects from a mere copy or photograph of the world. In the first place it is rather simplified. As we have seen, we select from the bewildering variety of impressions we receive, and our picture emphasizes those things which are of particular interest and ignores many which are of no interest. We can walk past a house every day for years, and so completely fail to take any notice of it that we have no idea what it looks like. Secondly, our picture of the world is greatly dependent on our previous experience. When I come first into a city my feeling is quite different to that which I have when I have 'found my bearings'. If I come up from an unfamilar underground station, sometimes for example I do not know whether I am facing north or south, and if by chance I get my bearings wrong it is quite an effort to get them turned round, so that my immediate picture fits in with the general picture of the town that I have previously made. The room in which I am writing has been familiar to me for years and my impressions of it are quite different to those I had when I first saw it. This means that the sense impressions of the moment are modified by and in some way cohere with the sense impressions of the past. My actual picture is therefore a composite one made up of both the actual sense impressions of the moment and my previous experience. It has a continuity with my own previous life. It is not merely that I recognize the room as the same one that I saw before; what happens is that the sense impressions of the moment are joined in some way with my previous pictures, so that the successive pictures

cohere into something which has continuity and thus provides an element of permanence to my life.

The first task in the development of an infant is to make this construction – a correspondence between the picture of the world and the world itself. In his first few months he gazes intently at bright objects and tries to touch them, feel them, suck them and so on. He is trying to make a coherent picture of his sense impressions. He is also trying to perform satisfying actions – often very simple ones like sucking at the breast or getting an object he wants, and he is thus learning at the same time to make muscular movements. It is probably very important that the construction of the picture of the world shall go on concurrently with the learning of muscular movements. Thus the child learns to make muscular movements within the framework of his picture of the world. Obviously this method of dealing with the world is possible only in a long infancy when the young animal is sheltered. The young animal is not initially 'teachable'. He constructs his own picture of the world and learns his own actions within it.

People do not merely use their limbs to perform actions; they often manipulate objects as extensions of themselves, as also can monkeys to a limited extent. Human beings use tools of many kinds, and manipulate them within their mental picture in much the same way as the limbs themselves. For example, a motor car is a tool for moving about, and when you drive it you precast its position in relation to other moving objects within your mental picture, and your life depends on the extreme accuracy of your forecasts.

This picture of the world is not a static one; it goes along with us throughout our lives, being constantly added to and enlarged. The new experiences are united in some way with the previous ones to make a coherent whole. How this is done is not known. It is obvious that my picture at the moment involves elements which come from both the present sensation and from past sensations. It involves memory and recognition. It also, as we have seen, involves a great degree of selection. This means that the human attitude to sense impressions is never a merely passive one. It involves the 'activity' of always endeavouring to interpret one's sense impressions, of trying to make sense of them and to fit them to previous experience. It is essential for a human being at all times actively to try to understand and make sense of the impressions he gets from the senses.

A consequence of this is curiosity. We notice particularly elements of the picture which do not fit in or are unusual in some way. What we see has to agree with what we hear, or we will try to find out why not. We are particularly interested in what is going on at the fringes of vision and often in what lies beyond the visible. To make a coherent picture of their experience is a deep necessity for human beings. If they are unable to do so they are unsatisfied and show signs of uneasiness and discomfort which they try to remove. The practical usefulness of this kind of curiosity is obvious. But it is also a habit which leads to wider and wider attempts to interpret experience and has perhaps been the driving force behind mankind's continued effort to understand itself and the universe.

REFERENCES

Lorenz, K. Z. 'The comparative method of studying innate behaviour patterns'. *Symposia of the Society for Experimental Biology, No. IV: Physiological Mechanisms in Animal Behaviour,* p. 221 (1950).

Tinbergen, N. 'The hierarchical organization of nervous mechanisms underlying instinctive behaviour'. *Ibid.*, p. 305.

Craik, K. J.W. *The Nature of Explanation.* Cambridge University Press, 1943.

COLOUR SPECIFICATION AND MEASUREMENT

C. L. BOLTZ

A PHYSICIST who without previous experience or warning found himself committed to work on colour and colour vision would soon be disturbed by a feeling of discomfort, even distrust. An introspectionist psychologist, committed to work with the physicist, would have the same feeling. The cause of this feeling would be the same for both men, namely, the *privacy of sensation.* Most physicists need not consider it. Psychologists are very familiar with it and many of them have no faith at all in attempts to measure it with physicists' tools.

This fact has made trouble in colour science for a very long time. One man looks at a coloured surface and says: 'Green'. Another man looks at the same surface and also says: 'Green'. Yet each man is having a unique experience, and the experiences may be quite different. Indeed in some cases we know they are, for colour-defectives exist, and it is not easy to discover their defects by attending to their verbal descriptions: special tests have to be devised to prove that a colour-defective person is abnormal. It is therefore understandable that people are puzzled to know how, and indeed *if,* a private experience can be subjected to specification and measurement.

Measurement is the use of a comparator, a 'yardstick', divided into convenient equal parts in terms of an agreed 'unit'; this much we all know. But this part of measurement is merely for arriving at a convenient and agreed scale of numbers. When we look further we discover that measurement, after all, is the production of a *negotiable entity* that can be submitted to the rules of mathematics. For the widest possible use in science and industry it must be agreed, standardized, reproduceable, invariable with respect to other recognized standards, and easy to use. Very few of our standards of measurement satisfy all these criteria; the standard yard varies with respect to the red wavelength of cadmium, and the standard of light intensity until a few years ago was variable and not easily reproduceable. Further, there can be no such thing as an 'exact' measurement; all that can be achieved is closer and closer approximation.

A specification of a colour is an agreed negotiable entity, which may, or may not, be capable of measurement. A word is, in itself, a negotiable entity, but of a kind which cannot be measured. For countless generations, English-speaking people have used the words 'red', 'blue', 'green', and so on, as negotiable entities for private sensations. These words have to be learned by experience and so there is no guarantee that the relationship between a word and the sensation it represents is the same for everybody. In fact for about 8 per cent of the population the relationship is so different from that of everybody else that the people in this group are said to be colour-defective. For the other 92 per cent, however, the words do bear a rough relationship to the sensations; this can be demonstrated by the amount of agreement between the performances of these 'normal' people when tested with apparatus. But there are not enough words in the language to cover the thousands of different colours seen in everyday life – D. B. Judd, of the National Bureau of Standards, has estimated that there are ten million discriminable colours. And words are of very little use when the people using them are without the coloured light or surface that is being reported on. This difficulty can be reduced by having a 'colour atlas' of specimens with their appropriate names; every paint-manufacturer's catalogue is such an atlas, limited though it is to the comparatively few stock paints being marketed. For really wide use a much greater subdivision than is possible with words is needed. This subdivision can be achieved with code letters or numbers – *R*1, *R*2, *R*3, etc., could be used for different specimens of red. But as soon as such a codification is attempted the codifier is compelled to systematize. He can do this by choosing specific mixtures of colour and letting the code show the proportions. This is how F. W. Ostwald started his colour system in 1917. The systematizer can choose variables corresponding, he believes, to private sensation-variables. This is how A. H. Munsell started in 1905. The variables that he used will be mentioned later. Alternatively, the systematizer can choose graded coloured glasses to provide filters for white light, number the glasses in a certain sequence, and rely on the user to make a visual match between the filtered (and therefore coloured) light and the colour he is trying to match. This is how J. W. Lovibond started in 1884 with his Tintometer.

These three named systems of colour specification are widely used to-day. They satisfy some of the criteria mentioned earlier; all of them are agreed (among their users), standardized, and reliable. Their codification is simple. The colour atlases are easy to use, but not all pigments are permanent and some are notoriously fugitive; so despite all the efforts of the atlas producers to use every available technical resource to ensure the permanence of their coloured specimens, they are still subject to change with time. The Tintometer coloured glasses, on the other hand, do not fade; they are permanent. But this system does not include an atlas; use of the specification depends on the possession by the client of the necessary colorimeter. The Ostwald system is also associated with some theories about colour 'harmony' and so has had considerable vogue among artist-technicians, and some artists' suppliers have produced paints specified according to the Ostwald system. The Munsell system was in the field very early in the United States and is there very widely used indeed. The Tintometer system allows of the specification of transparencies, especially solutions, as well as coloured surfaces. These facts have doubtless contributed considerably to the survival of the three systems and account for the great interest taken in them by colour scientists in the past decade.

In the past century and a half there have been several attempts to make other systems, but most of these have very quickly disappeared. New ones are still appearing but have a very stiff fight to establish themselves because of the firm commercial position of these three. One British system, that of the British Colour Council, has established itself among commercial users of colour in this country, but it is not yet clear what principles are relied on in its systematization. It is notable for having produced atlases showing the effect of surface texture on colour appearance and for the careful listing of names applied to colour materials. There is also the Maerz and Paul *Dictionary of Colour* of 1930, containing a comparative list of terms used by different systematizers. There is P. Ridgway's *Color Standards and Color Nomenclature* of 1912, originally devised for describing the plumage of birds. Moves have been started for the re-examination and revision of these two productions.

It may be asked whether the use of words alone can continue, even in biology. Admittedly some of the dictionaries have far more colour specimens than are in the Ostwald and Munsell atlases. Yet the intensive work done by colour physicists in the past twenty years has altered the whole position of colour specification and measurement. Colour provides common ground for biologist and psychologist and physicist if they use a system of colour specification and measurement expressed in terms that can be understood by all three. This is all the more important in view of the fact that the physicists' systems distinguish variables such as luminance and dominant wavelength. If this had been done by biologists and psychologists, some confusions about whether certain animal and human reactions were due to light as energy or specifically to colour quality would never have arisen.

None of the three systems so widely used really constitutes in itself an adequate system of measurement, though the codification lends itself to subdivision that is getting very near to something not easily distinguishable from measurement. The Munsell system in particular is based on private variables and the subdivision into steps thus provides graded series of negotiable entities that come very close to a 1 : 1 correspondence

with private sensations. All the same, none of the systems gives numbers that can be dealt with by the rules of mathematics.

For the master system of colour measurement there is another approach. The transfer from private sensation to a negotiable and measurable entity is achieved by the securing of agreement among normal observers under standard laboratory conditions. All normal observers agree not only that they can match any colour whatever with a second colour made by mixing together three monochromatic lights but that the quantities of these three stimuli for any one match are the same for all the observers, allowing for variability that can be dealt with by the usual statistical techniques. The three monochromatic lights are taken from the red, green and blue regions of the spectrum, and must be precisely specified. There is no one set of stimuli that must necessarily be employed: there is a freedom of choice, which fact has been used to meet practical convenience, in ways which will later appear. The quantities of the stimuli can be measured in absolute units (in ergs, for example) or relatively. The transfer from private sensation to a negotiable and measurable entity is thus made. The rest is manipulations by ordinary arithmetic and algebra, and by graphical methods, giving us the internationally recognized C.I.E. system of colour measurement and specification. It was established in 1931 by the Commission Internationale de L'Éclairage (C.I.E.). In the United States the initials are usually changed to I.C.I. (for International Commission on Illumination), a procedure that is not encouraged in Great Britain for obvious reasons.

Can it be said, then, that sensation has been measured? The question is seen to be irrelevant. What has been secured is agreement among normal observers about the matching of a test colour by a combination of standard colours, under conditions that eliminate many of the features that contribute to the perceptual knowledge of colour (such as simultaneous contrast, the scanning of a large patch by the eye, and so on). It is fair to assume that the two fields thus matched correspond with equivalent sensations, but what is equated is the stimulus of one field with the stimuli for the other. For practical purposes a measurable entity has been agreed. That is what matters. Can it be said that the C.I.E. system is fundamental, i.e. directly related to private variables revealed when a person judges a colour? At first sight, no. Nevertheless, objective variables that correspond with these private variables can be inferred from the C.I.E. specification.

PRIVATE (SUBJECTIVE) VARIABLES

All practical judgements of colour are made by human observers, whether in the judgement of colour films, interior decoration, pictures, biological specimens, or anything else exhibiting colour. This is a powerful argument in favour of visual colorimetry as against instrumental work in which the eye is replaced by a photoelectric cell. In all these visual judgements the private variables play an important part.

A normal observer, when presented with a large array of coloured patches and asked the right questions, is able to make three different series of discrimination. The classification by *hue* is obvious; it is the classification made by an observer when he says, 'This is blue; that is red.' He can also make a series of patches of the same hue but differing in *lightness* (or *luminosity* if he is judging coloured light instead of a coloured surface). It may be supposed that a certain amount of a chromatic pigment is mixed in the different patches with different amounts of a black pigment. Clearly, the greater the quantity of black, the less the proportion reflected of any incident light, though the amount of chromatic pigment is the same in each patch. This is the judgement called lightness. Yet a third series can be made of constant hue and lightness, but differing in steps of colourfulness: '*a* is a redder red than *b*'. The quality being judged, if an estimate of the proportion of chromatic colour present, is called *saturation*. But if the quality is an estimate of the actual amount of chromatic colour present, and not a proportion, then another word is needed. Outside the Munsell system there is no such word. Munsell calls it *chroma*. The difference can be seen if it is imagined that a quantity of chromatic pigment is mixed with increasing

amounts of white pigment. If the judgement made is that of a proportion, then clearly the proportion of chromatic pigment is decreasing, and so therefore is the saturation. Yet the quantity of chromatic pigment, and therefore the chroma, are still the same.

Not all these variables are independent, although different judgements can be made. In the same series of mixtures, lightness is increased as saturation is decreased, both these judgements being affected by the addition of white to the mixture. And if a second series is made, in which both lightness and hue are constant, then saturation and chroma are seen to be equivalent. In fact, there is sometimes debate as to whether the distinction between saturation and chroma is useful. There are also other possible private variables, which, again, are not independent of those more widely used. With suitable training an observer in the pigment industry can make judgements of the amounts of chromatic colour and of black and white in an observed colour; and workers in the dyeing industry can be trained to give still different judgements. However, outside particular industries, the judgements used for the creation of colour-series are hue, lightness and saturation (or else chroma), the three first being the variables commonly accepted in scientific work. Even these, however, are not wholly independent, for it is possible to make a series of the same chromatic pigment mixed only with black and white and to get a change in hue. Nevertheless, the changes for wide ranges of practical colours are slight, and the near-independence of the three variable can be considered as complete independence for many practical purposes.

Any series of, for example, changing lightness can be arranged in equal-appearing intervals, so that an observer can say: '*b* is as much lighter than *c*, as *a* is than *b*'. This fact is known in a number of other sensation phenomena as well as in those of colour. It is not known in what way the successive 'equal' steps are changing. Fechner said that sensation varied as the logarithm of the stimulus, so that stimuli increasing in geometrical progression gave rise to sensations increasing in arithmetical progression. There has been controversy for more than a century about the validity of his assumptions in the argument leading to his 'law'. But for the physicist engaged in trichromatic colour-matching, the truth or falsity of the Fechner Law is an irrelevant question: for the physicist produces a public scale of measurable stimuli that agrees with the operations of normal observers when faced with two fields to be matched for colour.

THE C.I.E. SYSTEM

The C.I.E. system is based on trichromatic matching. If the three monochromatic stimuli are *r*, *g* and *b* (standing for red, green, and blue), and the quantities of each required to match a colour *C* are *u*, *v and w* respectively, then an equation for the match can be written as follows:

$$C = uR + vG + wB$$

In this equation, uR means 'the quantity u of the stimulus R', and similarly for vG and wB.

If this equation is divided by $(u+v+w)$, we get:

$$\frac{C}{u+v+w} = \frac{u}{u+v+w}R + \frac{v}{u+v+w}G + \frac{w}{u+v+w}B$$

Obviously,

$$\frac{u}{u+v+w} + \frac{v}{u+v+w} + \frac{w}{u+v+w} = 1$$

So, if we write new symbols, u′, v′ w′ for the new coefficients *R*, *G* and *B*, we obtain the relationship that

$$u' + v' + w' = 1$$

Then, calling the quantity *C* unity, which can be formally justified, we get:

$$1C = u'B + v'G + w'B$$

This is a unit equation. In making it we have, so to speak, abstracted light as total energy and are left with an equation expressing colour *quality* according to the proportions of the

three stimuli needed. In other words it is a *chromaticity* equation. As the sum of the coefficients is unity, only two need be known and then the third can be found by simple arithmetic. This fact makes possible the use of a two-dimensional diagram on which only two of the coefficients are plotted. The result is a *chromaticity chart.*

Some twenty-three years ago, W. D. Wright, of the Imperial College of Science and Technology, London, and J. Guild, of the National Physical Laboratory, working independently, obtained careful figures for a number of normal observers doing trichromatic colour-matching. The matching was done with a monochromatic light as the test colour and was continued through the whole of the visible spectrum. The C.I.E. accepted the values to establish the 'standard observer.' It then went on to discuss the establishment of three fixed stimuli as the basis of the system: for there is a freedom of choice that can be exercised. There were disadvantages in accepting any one experimenter's stimuli, chiefly because it had been found that when an observer worked with such stimuli, there were some spectral test colours that could not be matched with positive quantities of the three matching stimuli: sometimes a 'negative' quantity was necessary, obtained in practice by adding a positive quantity to the test stimulus. In other words, one of the plus signs in the equation would have to be replaced by a minus sign. This necessity has been a common cause of misgiving, although it is of no further importance than to the extent that procedure in measurement is affected. However, since the colour-specifications obtained were likely to be used and quoted by relatively unskilled technicians, a system necessitating any negative values was considered unsuitable. It was therefore taken as one criterion in establishing the three C.I.E. stimuli that only positive signs should be needed. This could be achieved only by making the C.I.E. stimuli hypothetical, although they could be related to real stimuli, and *vice versa,* through a simple algebraic transformation. The slight inconvenience which results is thus confined to the relatively small number of colour physicists who make the measurements;

while the far greater number of technicians who use the specifications have the advantage of simplicity.

A second criterion came from the shape of the line that is obtained on the chart by joining up the positions corresponding with the pure spectral colours. It is known as the spectrum locus, and is seen in Fig. 1 as a line that is straight at the

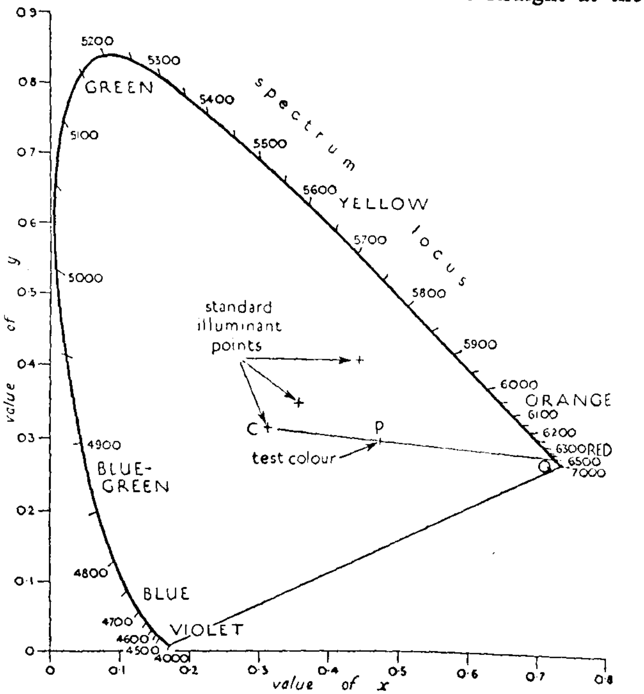


Fig. 1. A C.I.E. chromaticity chart [*x*, *y*] with the spectrum locus drawn. The point *P* represents a test-colour, whose *x*, *y*, *z* values are 0.48, 0.3, 0.22 respectively, with standard illuminant *C*. *Q* is the point at which *CP* cuts the spectrum locus on the side nearer to *P*, and gives the dominant wavelength of the test-colour (6,500 angstrom units in this case). The other two standard illuminant points correspond with standard illuminants *A* and *B*, not used in this measurement.

beginning of the red end of the spectrum, and remains straight up to the region of spectral yellow; but bends round in the section denoted as green, and thereafter descends to the region of blue and violet at the bottom-left of the chart. The straight section is still straight, irrespective of the particular choice of stimuli, but its position on the chart depends on the stimuli chosen. It is convenient in use that this straight section should coincide with a line at 45 degrees slope joining the value 1 of one stimulus to the value 1 of the other stimulus. This criterion, as well as the first, can be satisfied in a multitude of ways, merely by juggling with co-ordinate geometry. One more criterion could thus be applied. The one adopted arose from an ingenious suggestion made by D. B. Judd. He pointed out that, whereas the chromaticity chart showed only the relative values to give colour-quality, one variable in the actual matching was the total quantity of light energy – the luminance for a light source, or the luminance factor for a reflecting surface. So why not arrange the three proposed stimuli so that the value of one of them should give the luminance factor directly? This was accordingly done, and the three stimuli *X*, *Y*, *Z* were thus established, the corresponding chromaticity co-ordinates in the unit equation being written as *x*, *y* and *z*. The three stimuli are hypothetical in the sense that they cannot be realized with practical laboratory apparatus: the *X* stimulus is 'redder' than any attainable red, *Y* is a 'greener' than green, and *Z* a 'bluer' than blue.

Statement of the *X*, *Y*, *Z* values, and of the *x*, *y*, *z* values does not completely specify a colour: as everyone knows, the quality of the illumination influences the appearance of a surface colour. So three standard illuminants were established, *A*, *B*, *C*, corresponding roughly to artificial light by a tungsten-filament lamp, sunlight, and diffuse daylight respectively. These illuminants are specified by their colour temperature, i.e. the temperature of a black-body radiator that would give light of the colour indicated.

The *Y* value of the specification is the one that gives the luminance factor. So a complete C.I.E. specification needs to give only the Y value, the illuminant, and the three chromaticity co-ordinates x, y, z. An example is: $x=0.48$, $y=0.3$, $z=0.22$, $Y=0.09$, illuminant C. The point on the C.I.E. chromaticity chart representing this colour is shown as P in Fig. 1, C being the point representing the illuminant C and so the 'white point' for this specification.

The system can be used in an enormous number of ways. The establishment of a system of numerical values means that normal mathematical manipulation can be done. The commonest chromaticity chart is that giving the x and y values, but any two of the three chromaticity co-ordinates can be used when required. The chart can also be distorted to show characteristics needed for specific purposes. For example, the least perceptible difference of colour quality is represented on the normal chart by a different distance according to the hue, but by distorting the chart the distances can be made equal. The distorted chart is then much easier to use than the normal one for anyone whose chief concern is the least perceptible difference between one colour and another.

A C.I.E. specification is not readily recognizable, as it stands, in terms of colour appearance. The one given as an example can hardly be recognized from the figures as a sort of dull red. But a little work on the chromaticity chart gives variables more nearly related to private variables. Thus, the line CP can be drawn, and continued until it cuts the nearer side of the spectrum locus. This is shown at Q. The monochromatic light represented at this point, shown in terms of wavelength, then gives the colour's *dominant wavelength.* In the example illustrated, this is 6,500 angstrom units, well in the extreme red. Also, distance CP divided by the distance CQ, gives the *excitation purity,* in this case about 0.41. The luminance factor is already given in the specification. These three objective quantities – dominant wavelength, luminance factor, excitation purity – correspond closely to the private variables of hue, lightness, and saturation. There is not a 1 : 1 correspondence. For instance, a series of constant dominant wavelength and increasing excitation purity is not necessarily seen as one of

constant hue. The cause of this is that, as already stated, the private variables are not mutually independent. However, for many practical purposes the departure from a 1 : 1 correspondence is very small indeed and the three objective variables inferred from the C.I.E. data can be taken as representative of the private variables.

In twenty years this C.I.E. system has established itself throughout the world among colour scientists. It remains the master system. But it is not in itself associated with a colour atlas, and though as a specification its formulae can be readily used, for serious scientific work an elaborate colorimeter is needed. For the day-to-day user of colour in pigments a simpler type of specification is more suitable.

THE OSTWALD SYSTEM

F. W. Ostwald (1853–1932) was a famous German physical chemist. He realized that people using colour do not see it as a tiny field in a colorimeter. He was in revolt against the Newtonian analysis and preferred the Goethean. He also had ideas on colour 'harmony' for design. All these facts biased him in the design of his system. Accordingly he arranged his pigments of 'full colour' – the most saturated surface colour obtainable – round the circumference of a circle in such a way that those diametrically opposite were 'complementary'.

To produce other colours, Ostwald mixed his full colour with white and black. He took the Fechner Law as being correct and so chose his mixtures in quantities that increased in geometrical progression and thus obtained what he believed to be subjectively equal steps. His mixing was originally done on a rotating disc with sectors of full colour, black, and white. The component quantities, however mixed, were thus bound to add up to the same total quantity. Thus he obtained the relationship *white + black + colour* = 1. For any one hue he chose mixtures of white and black to give supposedly equal subjective steps and built the several mixtures into an equilateral triangle with full colour at one corner, white content at one corner, and black content at the third. The black-to-white base is placed vertical with white at the top. The full-colour corner therefore stands out half way down and fits into the colour circle of full colours, which is horizontal. Thus the complete array of specimens, the 'colour solid', makes a double cone, base to base, with vertical mid-sections consisting of two of his equilateral triangles. One such triangle is shown in Fig. 2. The axial base is an achromatic series from a grey consisting of 96.5 per cent black and 3.5 per cent white to a grey made of 11 per cent black and 89 per cent white. A series parallel to the *WC* base line has the same percentage of black

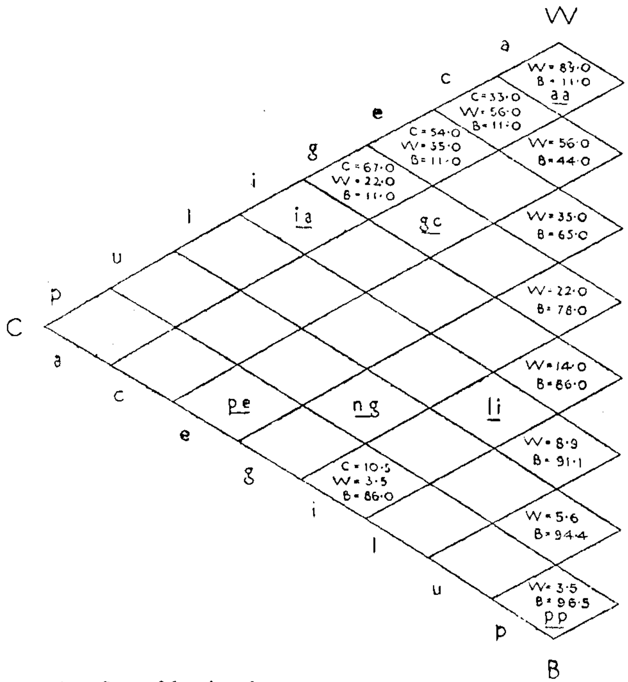


Fig. 2. An Ostwald triangle, one half of a vertical mid-section through the colour solid. Mixtures are shown for some of the spaces. The method of indicating the position in the triangle is also shown in several cases (here underlined for clarity).

in each specimen. A series parallel to the *BC* base line has the same percentage of white in each specimen. Anyone who cares to do the arithmetic can verify that the figures for the percentage of white in the vertical achromatic series are so arranged that ratios of successive pairs all the way up are fairly constant. Thus 5.6 divided by 3.5 gives almost the same answer as 8.9 divided by 5.6, and so on. The logarithms of these numbers thus make nearly an arithmetical progression.

The codification is done by giving a number to each hue, i.e. each full-colour or colour of a certain dominant wavelength. The white content is given by a letter, the eight letters – *a, c. e. g, i, l, n, p* – being used in that order to represent decreasing amounts of white content. The same letters are used in reverse order to represent the black content, so that *a* represents the minimum black and *p* the maximum. The specification is then written thus: *6ne,* meaning hue 6, white content *n*, black content *e*. The most recent atlas, produced in the U.S.A. in 1942 and called the *Color Harmony Manual*, has 943 specimens. The C.I.E. specifications of the mat specimens have been found experimentally by American colour scientists.

THE MUNSELL SYSTEM

A. H. Munsell (1858–1918) was an art teacher in Boston, Massachusetts. He therefore arrived at subjective variables first and then applied his knowledge of science to systematize. The variables used are hue, value, chroma. Value corresponds to lightness; chroma has already been described in a previous section. The Munsell colour solid is built from a central vertical trunk made of 9 subjectively equal steps of achromatic colour (grey) from a dark grey at the bottom to a very light grey at the top. Chroma is taken in horizontal steps from each of the trunk patches and these chroma patches extend horizontally. Hues are arranged in a horizontal circle, the most saturated being on the circumference. The colour solid has thus an irregular profile. A vertical mid-section consists of two groups, one on each side of the trunk, each group consisting of patches of constant hue. One of such groups is shown in Fig. 3. The number of subjectively equal steps of increasing chroma is different for different lightnesses. The section is for the hue yellow (Y) and the chroma steps starting from value 8 are seen to be seven whereas the chroma steps from value 5 are only four. These chroma steps are numbered 0, 2, 4, 6, etc.

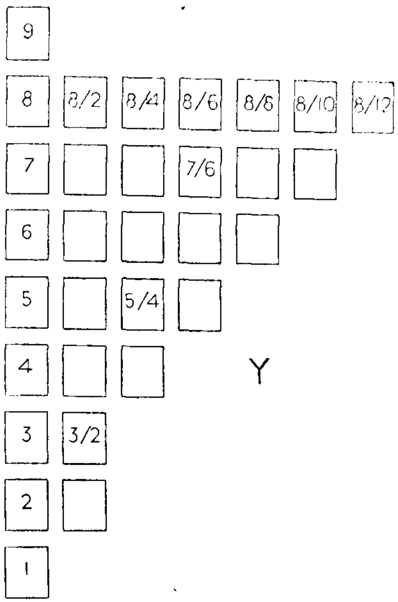


Fig. 3. Half a vertical mid-section through a Munsell colour-solid The hue is Y (yellow). The figures up the left side show value in the achromatic series. Pairs of figures, of which some examples are given (e.g. 8/2) show value and chroma.

A Munsell specification consists of a number, followed by one or two letters, followed by a second number, and then an oblique stroke followed by a third number; for example, 5PB3/12. The hue is given by the letter, and the division of the hue by the number preceding the letter. In the example, PB means purple-blue, and 5PB means 'hue 5 of purple-blue'. The second number gives the 'value', 3 in this case; and the

number after the stroke, the chroma. The most recent atlas consists of 960 specimens. It is called *The Munsell Book of Color* and was published in 1942. C.I.E. specifications of the Munsell specimens have been measured by American colour scientists.

COMMENTS ON THE MUNSELL AND OSTWALD SYSTEMS

Anyone who needs only a convenient atlas of coloured specimens will find nothing to choose between the two systems. Each atlas is beautifully produced, is expensive, and suffers from the impermanence of pigments. Each has been modified from the original and thus is not as restricted as the rigid systematics might suggest. For instance, the Ostwald colours are no longer produced by means of a colour disc. For practical purposes the Ostwald system has one great disadvantage. Because of the method of making a closed triangle for each hue, the most saturated colour is that obtainable at the time of manufacture. As new pigments are discovered of greater saturation they cannot be given an Ostwald notation without a complete overhaul of the system. Such new pigments, for example, have been produced recently in the daylight-fluorescent poster pigments now familiar to all town dwellers. The Munsell system has no such limitation and any more saturated pigment can be given a Munsell notation with a bigger chroma number.

Neither system is as good as its inventors hoped it would be. Colours of constant Munsell chroma, for instance, when plotted on a C.I.E. chromaticity chart show sudden irregularities, whereas mathematical analysis shows that the curves should be smooth. Earlier Ostwald colours showed even worse jumps. From the smoothed curves of the C.I.E. chart new notations can be obtained. This has been done by American colour scientists for the Munsell system and the re-notations have been published.

Anyone owning a new Munsell atlas and the corresponding C.I.E. specifications has a very convenient reference library. The fact that the Munsell system lends itself to extension as more saturated colours become available and the fact that the variables used are the private ones combine on the whole to make colour scientists prefer the Munsell system to the Ostwald. Practical users of colour such as architects and colour-consultants have testified publicly to the ease with which they can use the system, the terms hue, value and chroma becoming easily negotiable entities for employees without scientific or technical training.

THE TINTOMETER SYSTEM

This is different from the previous systems in that the user has to perform simple colorimetry. It was devised by J. W. Lovibond (1833–1918) of Salisbury in 1884. It provides a simple system of specification that gives some idea of what a colour looks like, and is independent of pigment-permanence. It can be applied both to surface colours and to the colours of transparent media. Further, in its most recent form it can be used to measure the C.I.E. specifications of a wide range of practical colours: all, in fact, except the most intense greens and blue-greens. This combination of qualities make the system very suitable for serious work where technicians with a little amount of skill are available, and where education in colour science is carried on. The colorimetry is visual.

Looking through the optical eyepiece of the colorimeter, the observer sees two fields side by side. One is the test field and consists of light that has passed through the transparent material being tested or else has been reflected from the opaque material being tested. The other field consists of light that has passed through subtractive transparent slides. There are three ranges of these slides, one red, one yellow, one blue. The slides are of coloured glass, are manufactured by the producing company itself, and are quite permanent. They are graded in such a way that their numbers can be added. Thus two yellow glasses of 0.5 and 2.0 can be added to make the same colour as a yellow glass of 2.5. The specification is merely the statement of the colours and the slide value of each, thus: 3.0 yellow, 2.5 red, 1.0 blue. This is an arbitrary code but is a

negotiable entity like that of the Munsell and Ostwald systems. The grading of the glasses is also arranged so that three of them of the same number, one from each range, together produce an achromatic colour or grey. The observer adjusts the glasses to get a match with the test field and then reads the answer from figures marked on the racks holding the slides. The instrument, which is much less expensive than the more elaborate sort of colorimeter used in the standardizing laboratories and the leading universities, is usable with all of the most common practical colours, and is especially good for those only slightly saturated, such as 'off-whites', which are found in many businesses concerned with papers and distempers.

The adaptation for the measurement of the C.I.E. specifications was introduced by R. K. Schofield in 1939. It depends on the fact that only two chromaticity co-ordinates are needed to specify colour-quality, the luminance factor being a separate variable. The match in the Lovibond-Schofield Tintometer is therefore made with only two sets of slides because the addition of the third set would affect only the luminance factor. To find the latter, use is made of a vane, known as the 'obturator vane', or more commonly as the Rothamsted Device, to block out part of the white light; the change in the total amount of light being shown as a dial-reading. There is a small added complication from the loss of light by reflexion when passing through the Tintometer slides. For this reason, compensating slides of plain glass are introduced on the test side of the instrument. Cells are available to adapt for the three standard illuminants of the C.I.E. system; and the observer is left effectively with two Tintometer slide values and a scale-reading. C.I.E. specifications are obtained from the experimental data by a simple form of manual convertor. For this purpose, use is made of a C.I.E. chromaticity chart for the x and z values, fitted with special cursors. In this way, the complete C.I.E. specification, and with it the dominant wavelength, and the saturation or the chroma, can be readily found.

To sum up, although experienced colour is a private sensation, a negotiable entity representing the colour can be obtained. When this is of a kind that can be subjected to the rules of mathematics, a system of colour-measurement has been achieved. The four systems of specification most widely used are those of the C.I.E., Ostwald, Munsell, and the Tintometer. Of these, the C.I.E. system is the only one that is to be considered truly a system of measurement. It is accordingly the master system, although in detail it has not escaped criticism, and long-term researches are now in progress to re-assess the curves accredited to the 'standard observer' from which the system starts. The Tintometer system gives both a specification and a method of measurement on the C.I.E. system, by means of simple and inexpensive visual colorimetry.

THE ACCELERATION OF CHARGED PARTICLES TO HIGH ENERGIES

L. L. GREEN

In order to penetrate the strong repulsive electric fields that surround atomic nuclei, nuclear physics has always required atomic nuclei travelling at high speeds as its main research tool. The first nuclear disintegration was observed by Rutherford using the high-speed nuclei emitted in the decay of radioactive materials. For studies of nuclear transmutations, there is a requirement for more copious supplies of these high-speed particles, at energies which can be varied. For this purpose various types of apparatus for the acceleration of particles were developed in the 1930's. The first nuclear transmutations with laboratory-accelerated particles were produced by J. D. (now Sir John) Cockcroft and E. T. S. Walton with an apparatus which accelerated charged hydrogen nuclei with a high voltage produced by the use of transformers, condensers, and rectifying valves. High voltages needed for the acceleration of particles have also been produced by electrostatic machines of the Van de Graaff type, which are similar in principle to the Wimshurst machine, so often used in laboratory demonstrations involving the generation of static electricity. The ultimate voltage obtainable with such devices is limited by the electrical breakdown strength of the air or other gas which surrounds the machine. Voltages of up to 5 million volts have been obtained with Van de Graaff generators surrounded by gas at high pressures. Work is in progress in the United States on two electrostatic generators designed to produce up to 12 million volts. These machines have the great advantage that particles accelerated by them can be made to have a very precisely determined energy which is of great use in the study of nuclear energy-levels.

Particle-energies in excess of these figures are needed for the study of interactions between fundamental particles. For the attainment of these higher energies it is necessary to use machines which rely, not on the production of high voltages directly, but on the use of only moderate voltages applied to the particle a large number of times, so that by repeated small accelerations a high ultimate energy is produced. In this case, the ultimate energy attainable is limited only by the number of accelerations which can be applied.

An instrument known as a linear accelerator, built by D. H. Sloan and E. O. Lawrence in Berkeley, California, was the first of these resonance accelerators. The instrument was used only for the acceleration of heavy ions which were of no use for the study of nuclear transformations, but the machine was the forerunner of the cyclotron and so is worth consideration.

The linear accelerator consisted essentially of a series of hollow electrodes arranged in a straight line. Alternate members of the series were connected together and an alternating voltage supply connected to the two sets of cylinders. At any

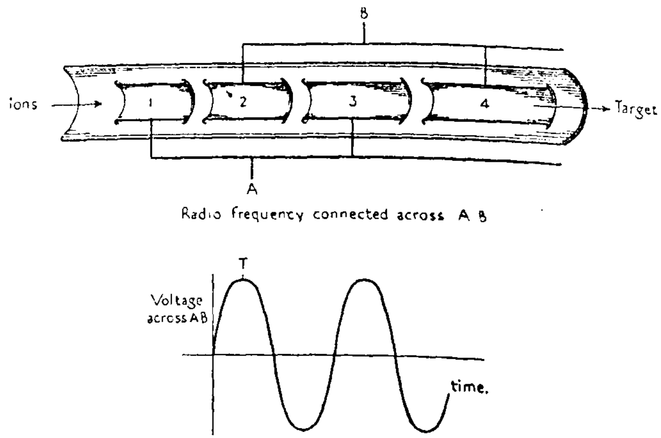


Fig. 1. To show the principle of the linear accelerator (see text).

instant there is therefore a voltage between alternate cylinders which will be varying in the manner illustrated in Fig. 1. At point *T* in the cycle all the even cylinders will be positive and all the odd-numbered ones negative. Half a cycle later the odd-numbered cylinders will be positive with respect to the even ones.

Suppose some positively charged ions now start at cylinder 1 and move from left to right down the axis of the cylinders. In moving through the first cylinder, they receive no acceleration. On moving into the region between cylinders 1 and 2, where there is a potential difference between the cylinders, the positively charged ions will be accelerated if the first cylinder is positive and the second negative. Whilst moving through the second cylinder the particles move with a uniform speed but at a higher speed than that at which they traversed the first cylinder. If the time taken by the ions to travel through the second cylinder is just equal to a half cycle of the alternating voltage supply, then when the particles reach the region between the second and third cylinders, the second will now be positive and the third negative, so that the particles will again be accelerated across the gap. If the cylinders are made progressively longer so that in spite of their increasing speed the particles take the same time to traverse each cylinder the particles will continue to be accelerated across successive gaps. The particles will emerge from the tube with an energy equal to the voltage applied across the gaps multiplied by the number of gaps traversed. With the highest oscillation frequency available in the 1930's the cylinder lengths required for protons were inconveniently long and it was necessary to work with comparatively heavy, slow-moving ions. Mercury and lithium ions were accelerated by Sloan and Lawrence.

The very much higher frequency oscillators developed during the war have enabled L. W. Alvarez to build a proton linear-accelerator of essentially this type at Berkeley. The oscillation frequency used is 200 megacycles per second and the protons are accelerated to 32 million electron-volts (MeV) in a distance of 40 ft.

THE CYCLOTRON

The cyclotron derives directly from the linear accelerator. The great length and numerous electrodes of the linear accelerator can be avoided if the particles are made to move on a spiral path by a magnetic field. In a uniform magnetic field a particle moving with a constant speed will move on a circle whose radius depends on the speed of the particle. A particle whose speed is increasing will move on a spiral of increasing radius. Provided the mass of the particle remains constant the time taken by a particle to complete one revolution remains constant, and is independent of the linear velocity of the particle. This is the principle which is used in the cyclotron (Figs. 2 and 3). The cyclotron was invented by E. O. Lawrence who built the first machine in 1931.

In its usual form the cyclotron consists of two flat, hollow semi-circular electrodes situated between the poles of an electromagnet. These electrodes resemble the two halves of a flat pill box slightly separated. Because of their shape they are usually referred to as the 'dees'. The dees are supported inside a hollow box, usually called the tank, by means of rods which themselves enter the box on insulated supports and may be connected in turn to a supply of radio-frequency voltage. The box is vacuum tight and continuously evacuated with high speed pumps, so that a vacuum of the order of 10^{-5} mm. of mercury is maintained. This box is placed between the poles of the electromagnet in the uniform magnetic field. The sides of the box have numerous ports through which various auxiliary pieces of apparatus may be introduced into the chamber through vacuum seals. The large-scale vacuum plumbing required by cyclotrons has been responsible for many developments in this technique.

In the centre of the box and placed symmetrically between the two dees is a source of ions. This usually consists of a low voltage arc in hydrogen, deuterium (hydrogen-2) or helium, according as we wish to accelerate protons, deuterons or alpha-particles. If, when a positive ion is produced at the centre of the dees, a voltage exists across the dees, the ion will be

attracted into the negative dee. It will acquire a velocity and will move in a circular path inside the negative dee. The region inside the dees is free from electric fields and the particles come under the action of electric forces only in the region between

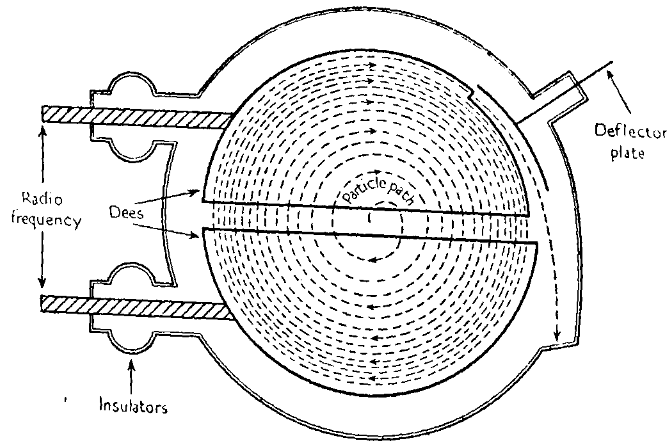


Fig. 2. To show the principle of the cyclotron: simplified plan view. The particles are accelerated at each crossing of the gap between the dees.

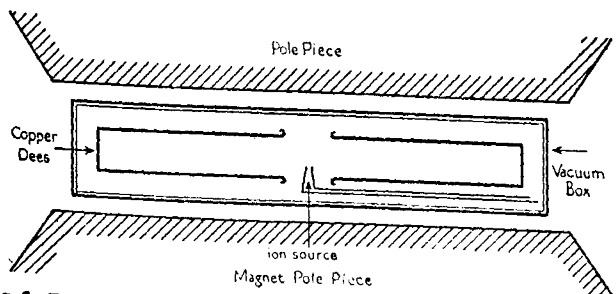


Fig. 3. To show the principle of the cyclotron: simplified vertical section through magnet pole-pieces, vacuum box, and dees. The particles are constrained by the magnetic field to follow a circular path (Fig. 2) at each stage of acceleration.

the dees. If the time taken by the particle for the half revolution is exactly a half period of the alternating voltage on the dees, on arriving again at the gap between the dees the dee voltage will have reversed and the particle will again be accelerated across the gap. The particle will thus move across into the opposite dee moving on a semi-circle of slightly larger diameter. If the oscillation frequency is adjusted so that the particle is accelerated across the gap on each of its first two traverses, because the time for each half-revolution is constant, the particle will be accelerated on each of its subsequent traverses of the gap. Thus the particle spirals out to the circumference of the dees, receiving a small impetus each time it crosses the gap, the radius of its path and its linear velocity gradually increasing. The ion source at the centre of the tank supplies a steady stream of ions which are accelerated in this manner. The radio-frequency period must be correctly adjusted for the mass of the ions and the value of the magnetic field, and only ions of the correct mass will remain in phase and will continue to be accelerated. The frequencies required for the oscillating voltage are in the frequency-range of ordinary short-wave radio.

On reaching the circumference of the dees the ions are brought out of the chamber by means of a deflector plate. A slot is cut in the side of one of the dees and a flat metal plate placed parallel to the dee circumference. This plate is charged to a high negative voltage and so attracts the positively charged particles which are thus pulled out of the dees. The tank is designed to allow the emergent beam to enter a target chamber.

The particles, in reaching the circumference of the dees, will have made many revolutions and travelled a path of many yards. In such a path length small inhomogeneities of the magnetic field and scattering of the particles by residual gas in the tank will cause the beam to diverge; unless some restoring force is provided the particles will hit the top or bottom of the dees and be lost. Unless there are focusing forces which confine the particles to the meridian plane between the upper and lower dee surfaces most of the particles will be lost before

reaching the dee perimeter. Fortunately in the central region of the dees the electrostatic forces between the dees have a focusing effect on the particles crossing the gap; and so also, at the edge of the dees, has the slight falling off of the magnetic field. These two effects tend to confine the beam to the meridian plane.

The maximum energy obtainable from a cyclotron is fixed by the diameter of the dees and the value of the magnetic field. The first cyclotron built by Lawrence had a magnet with pole faces 11 inches in diameter and produced protons of 1.2 MeV energy. Subsequently, machines with pole-faces 37-inch and 60-inch diameter were built. There are now many cyclotrons in existence, beam intensities varying from 10 to 500 microamperes being obtained. In this country machines with 37-inch pole-diameters are in operation in Cambridge and Liverpool, producing 4 MeV protons and 8 MeV deuterons, and at Birmingham there is a 60-inch cyclotron producing 20 MeV deuterons. Photographs of the Birmingham 60-inch cyclotron are reproduced as Inset 1 and 2, and of the Liverpool 37-inch cyclotron as Inset 4.

The cyclotron depends for its operation on the angular velocity of the particle remaining constant in the uniform magnetic field. The angular velocity of the particles depends on the mass of the particle and is constant only if the mass of the particles is constant. As predicted by Relativity Theory, the mass of the particle is not constant, but varies with its velocity. The effect is important only at very high speeds approaching the velocity of light; but in a 60-inch cyclotron, as a consequence of the high velocity, the mass of the emerging particles has already changed to 1.01 of the rest-mass. This means that at the edge of the dees the period of rotation of the particle in the magnetic field is larger than at the beginning of its acceleration. Because the period of the oscillating voltage applied to the dees is constant the particles after each half turn will arrive at a slightly later point in the radio-frequency cycle. After many revolutions they may arrive so late at the dee-gap that the voltage across the gap has changed to a decelerating voltage. This effect sets a limit to the energy to which particles can be accelerated with conventional cyclotrons, and the 60-inch cyclotrons are probably near the limit for economical operation for fixed-frequency cyclotrons. Although the energy obtained from such machines is ample for most experiments on nuclear disintegrations, the study of interactions between fundamental particles requires energies much in excess of the 20 MeV or so thus obtained.

Variations in the rotation-period of the particle with energy could be overcome to some extent by having a very high voltage on the dees so that the particle made only a few revolutions. It was originally planned to operate the 184-inch Berkeley cyclotron in this way with about 1 million volts on the dees. Two other methods of compensation for the increase in mass are possible. The magnetic field could be increased towards the perimeter of the magnet. This unfortunately also introduces forces on the beam which tend to make the particles diverge from the meridian plane, so that the beam is defocused. The other possibility is to leave the magnetic field unchanged and decrease the frequency of the oscillating voltage applied to the dees as the mass of the particle increases. Such a modulation of the frequency of the dee-voltage would be quite impossible with the conventional cyclotron orbit since the modulation frequency required would be much too high. Fortunately there is an effect which depends on this relativistic increase in mass which was pointed out independently by E. M. McMillan and V. Veksler in 1945 and which solves the problem. Essentially the same idea had been put forward by M. L. Oliphant in 1943.

SYNCHRO-CYCLOTRONS

If we consider the acceleration of the particle in the conventional cyclotron we find that the normal orbit is not phase-stable: there are no forces which tend to maintain the particle on an orbit crossing the dee-gap at the correct point on the radio-frequency cycle. For a particle moving in the region in which the mass of the particle is changing with velocity – the relativistic region – there is, in a constant magnetic field and

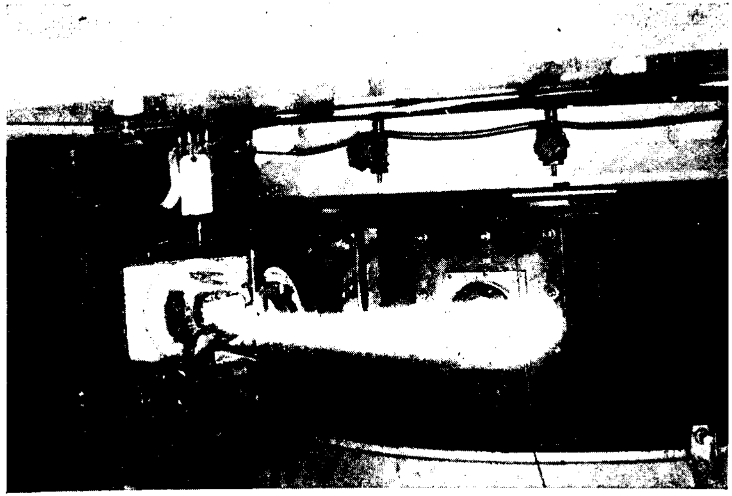
with a constant frequency, one energy for which the particle will move on a stable orbit. This will be the energy for which the angular frequency of the particle is equal to the angular frequency of the dee voltage. If the particle crosses the dee-gap when the radio-frequency voltage is zero, and about to become decelerating, the orbit will be stable and the particle will continue to move on it indefinitely. It is easy to see why this is so. If for some reason a particle had an excess energy and moved on an orbit corresponding to a slightly higher energy, the relativistic increase in mass would make the time of rotation slower, so that on its next arrival at the dee-gap, the dee-voltage will be slightly decelerating so that the energy of the particle will be reduced. This will continue until the particle eventually crosses the gap at zero dee-voltage. If the particle had too low an energy the reverse would happen and the particle by arriving at the gap slightly early would be accelerated until it arrived at the dee-gap at the correct time. So that, in this orbit, crossing the dee-gap just before the dee-voltage becomes decelerating, any particle getting slightly out of phase with the dee-voltage is pulled back into the correct phase. In practice the particle will execute oscillations around the phase-stable orbit, circulating on or around the orbit indefinitely, neither gaining nor losing energy.

Such a stable orbit can be made to expand, and the energy of the particle increased, either by increasing the magnetic field or by decreasing the frequency. If the frequency is reduced slightly the particles will next cross the gap in the accelerating region of the radio-frequency cycle and will continue to do this until the increase in mass is sufficient to bring them once more into resonance with the new frequency. This will be at a higher energy on an orbit of slightly larger diameter. If the frequency is changed slowly compared with the time of rotation of the particles, phase-stability is maintained and the particles settle down in the new orbit. The energy of the particles in a phase-stable orbit can also be increased by increasing the magnetic field and varying the frequency so as to maintain the particles in an orbit of constant radius. By these methods of varying the frequency or the magnetic field, the energy of the particles can be slowly increased up to the maximum. This maximum is set by the maximum magnetic field and the diameter of the largest orbit that is feasible.

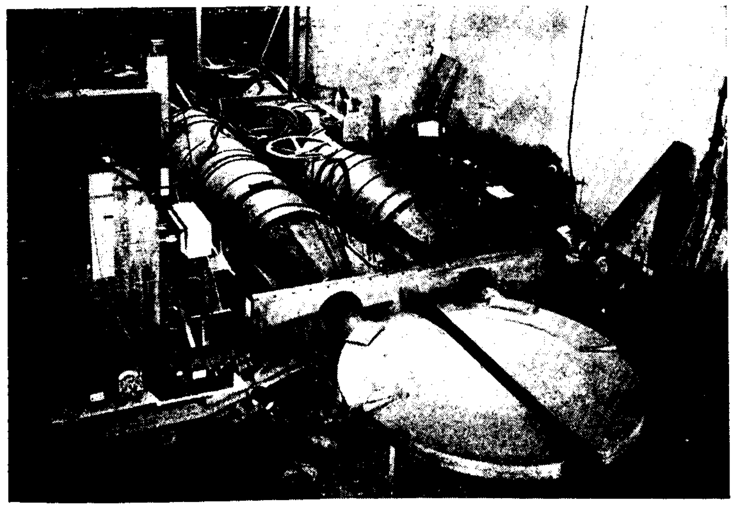
In practice both these methods have been employed. In large cyclotrons where the magnetic field is kept constant and the frequency varied, the instrument is referred to as a frequency-modulated cyclotron or synchro-cyclotron. In order to try out McMillan's idea before restarting work on the Berkeley 184-inch cyclotron after the war, the 37-inch cyclotron was converted for frequency-modulation. To simulate the conditions which would be experienced in the 184-inch machine by the relativistic increase in mass, the magnetic field was made to fall off at the perimeter so that a 13 per cent decrease in the frequency between the ion source and dee perimeter was required. The application of frequency modulation was immediately successful. Particles were accelerated in the 37-inch cyclotron to the full energy over the whole range of frequency modulation which was tried. The ion stability was so great that some of the particles were accelerated if the rotating condenser, which varied the frequency, was turned slowly by hand. In this case the ions would have revolved millions of times before reaching their final energy.

The frequency-modulation method was applied to the 184-inch cyclotron (Inset 3). The magnetic field is not strictly uniform in this, but is allowed to fall off slightly from the centre to the edges to provide the necessary field gradient for focusing the particles in the meridian plane. This means that the frequency must vary slightly more than is required to compensate for the increase in mass of the particle.

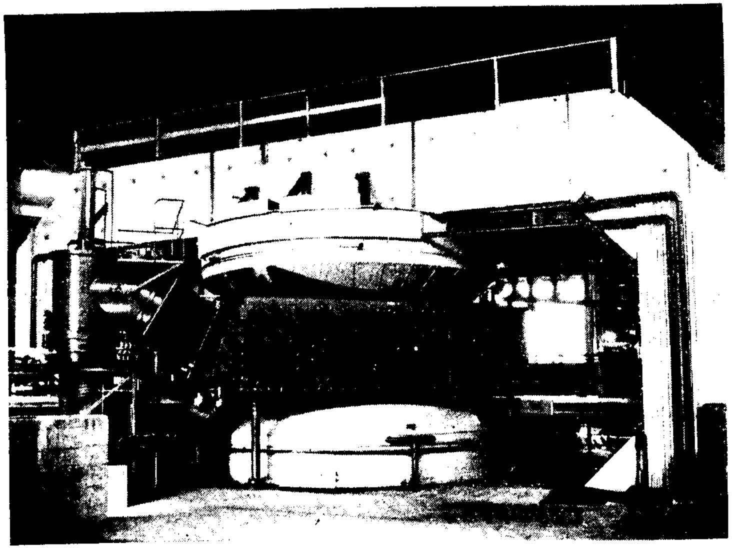
The use of the phase stable orbit leads to many advantages in the construction of large cyclotrons. Because the particle can now make a large number of revolutions and still remain in phase there is no need to have large voltages on the dees. This means that the radio-frequency power required is much smaller, and the engineering of the radio-frequency supply is therefore much simplified. The 184-inch cyclotron at Berkeley, for in-



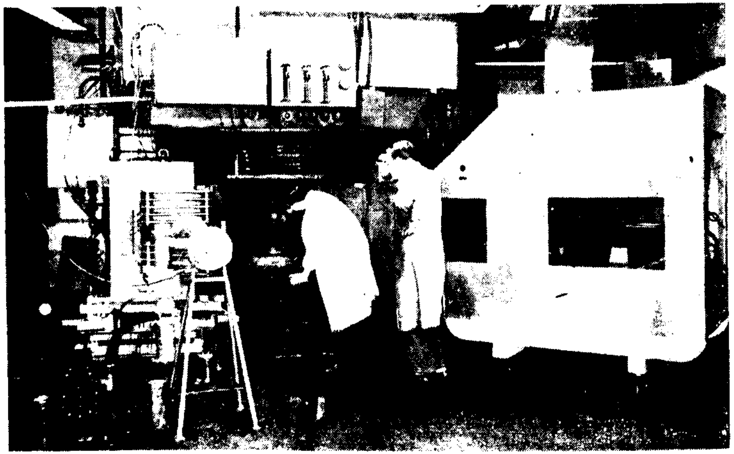
1. Birmingham University, 60-inch cyclotron: emergence of proton beam. *(Photograph by Dr W. I. B. Smith.)*



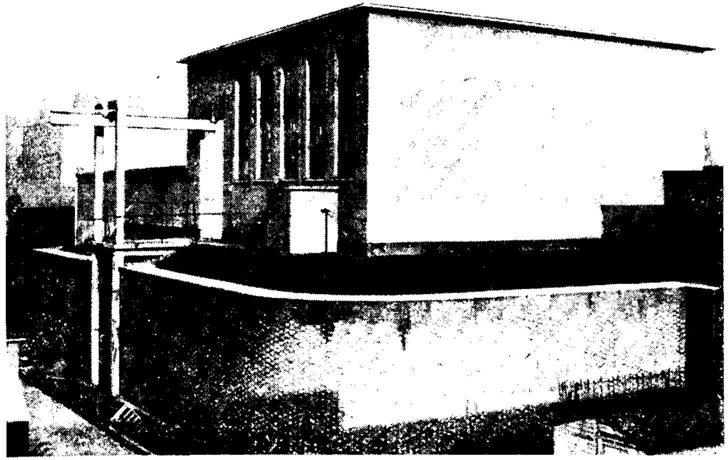
2. Birmingham University, 60-inch cyclotron: dees and transmission lines. *(Photograph by Dr D. M. Millest; courtesy Physical Society.)*



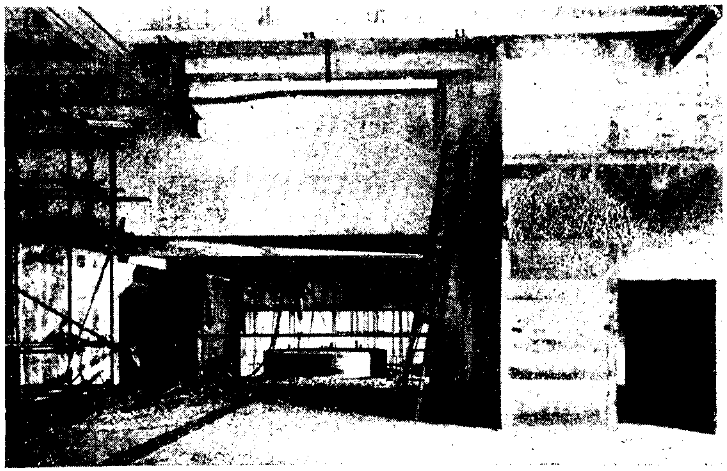
3. University of California, 184-inch cyclotron, the biggest of three machines built by Professor E. O. Lawrence. It was with this machine that meson particles (*Science News 16*) were first produced artificially. *(By courtesy of the University of California.)*



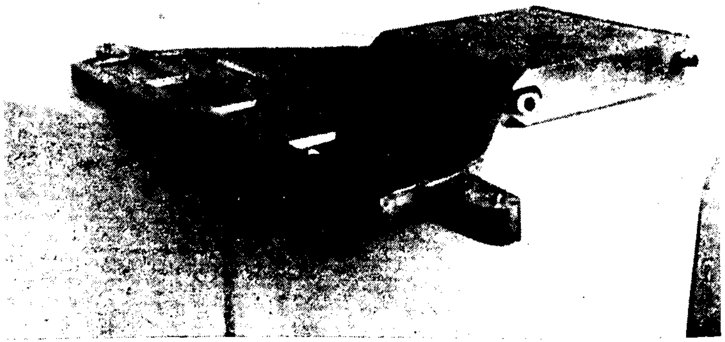
4. Liverpool University, 37-inch cyclotron. Small magnet, used to re-focus the emergent beam, can be seen on the left.



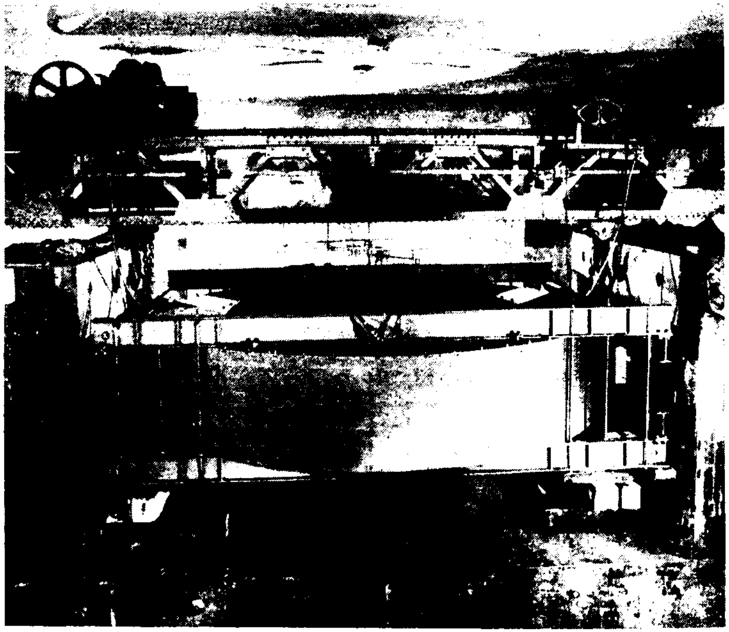
5. Liverpool University, 156-inch synchrocyclotron: external view of building. The cyclotron is below ground-level, and is built into the side of a hill. The cyclotron itself is surrounded by 6 ft of concrete and the rubble which has been excavated for the site.



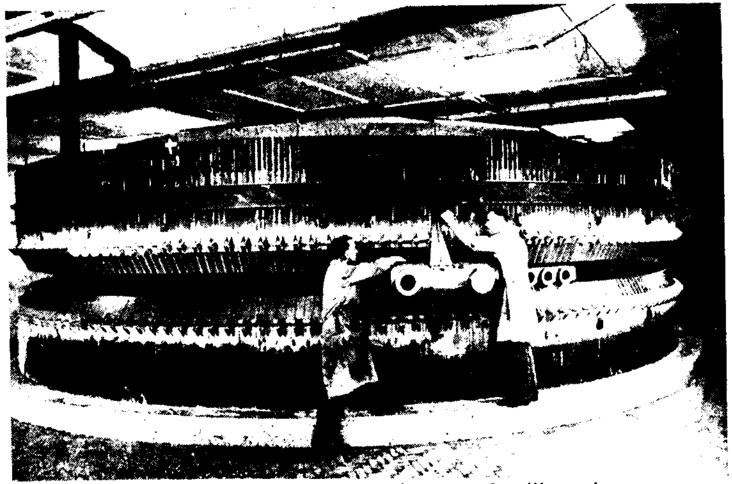
6. Liverpool University, 156-inch synchrocyclotron: general view of cyclotron room. The beam will emerge towards camera, and the large gap in the concrete shield will be filled with concrete blocks.



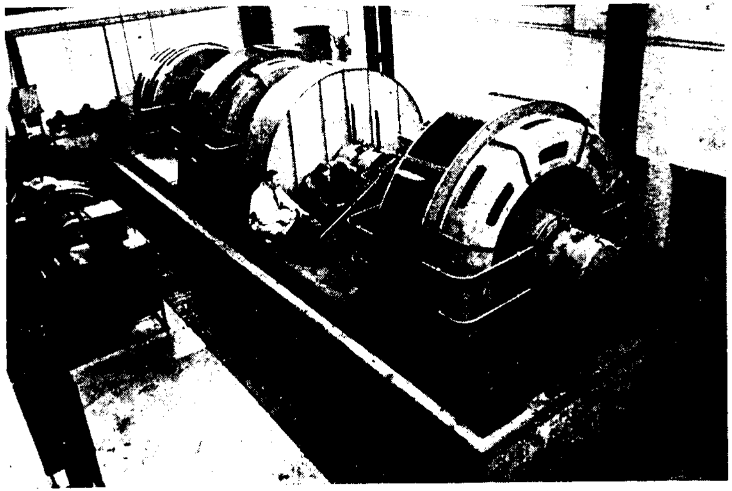
7. Liverpool University, 156-inch synchrocyclotron: dee of $\frac{1}{4}$-scale model.



8. Liverpool University, 156-inch synchrocyclotron: exciting coil being lowered into position on the lower pole-piece.

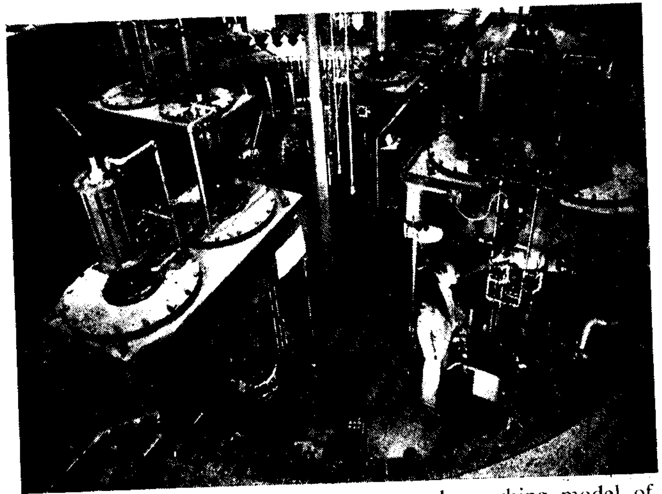


9. Birmingham University, proton synchrotron. It will accelerate protons to an energy of 1,300 million electron-volts.

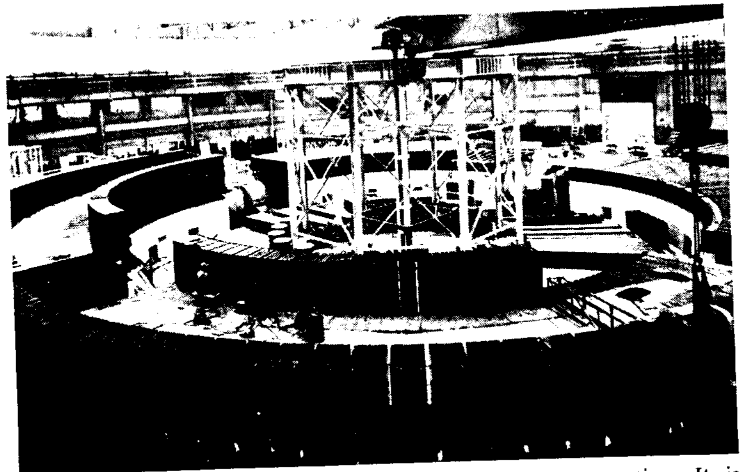


10. Birmingham University, proton synchrotron: the power system. From left to right: motor, direct-current generator, 37-ton fly-wheel, and second generator. The generators provide a peak current of 12,000 A. The peak voltage is 1,100 V. The fly-wheel serves both as a reservoir of energy and for its recovery.

11. Birmingham University, proton synchrotron: showing the laminated construction of the magnet. There are 2,400 laminations, each ½-inch thick, welded together. *(This and the previous two photographs are from the Central Office of Information, Crown Copyright reserved.)*



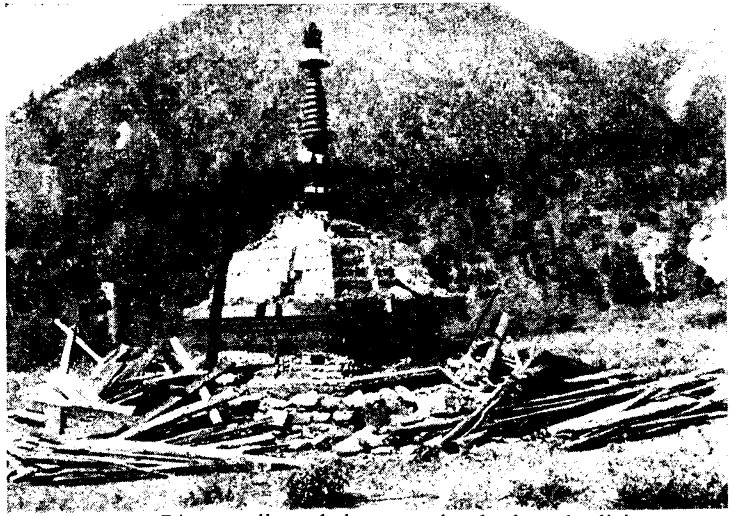
12. University of California, quarter-scale working model of Bevatron. The quarter-scale model was the first proton synchrotron to work. *(University of California, courtesy 'Scientific American'.)*



13. University of California, the Bevatron under construction. It is designed to accelerate protons to energies of 3,500 million electron-volts in the first place: later, it will be stepped up to 6,000 million electron-volts. *(University of California, courtesy 'Scientific American'.)*

14. Scarred mountainside near Rima, where Mr Kingdon-Ward was in camp on the evening of the Assam earthquake, 15 August 1950.

15. The La Ti torrent: 'The previous day the sparkling water revealed every boulder. Now it was dark coffee-coloured, but so thick with foam that it looked like café-au-lait'.



16. Damage at Rima: collapsed chorten, a bottle-shaped religious structure stripped down to the central pillar.

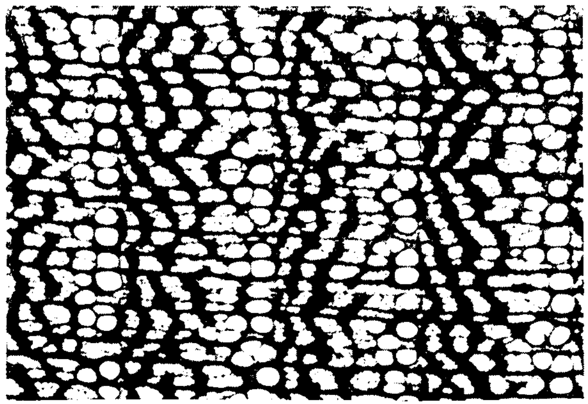
17. Damage at Rima: remains of monastery. The stone platform collapsed and the building canted over.

The structure and evolution of the tropical hurricane are discussed by R. W. James in this issue. Like a hurricane, a tornado consists of a rapidly revolving column of air: it extends downwards from a cloud, usually associated with a violent thunderstorm. Inset 18 and 19, below, show damage from the Buckinghamshire tornado on 21 May 1950. They are reproduced by courtesy of the *Bucks Advertiser*.

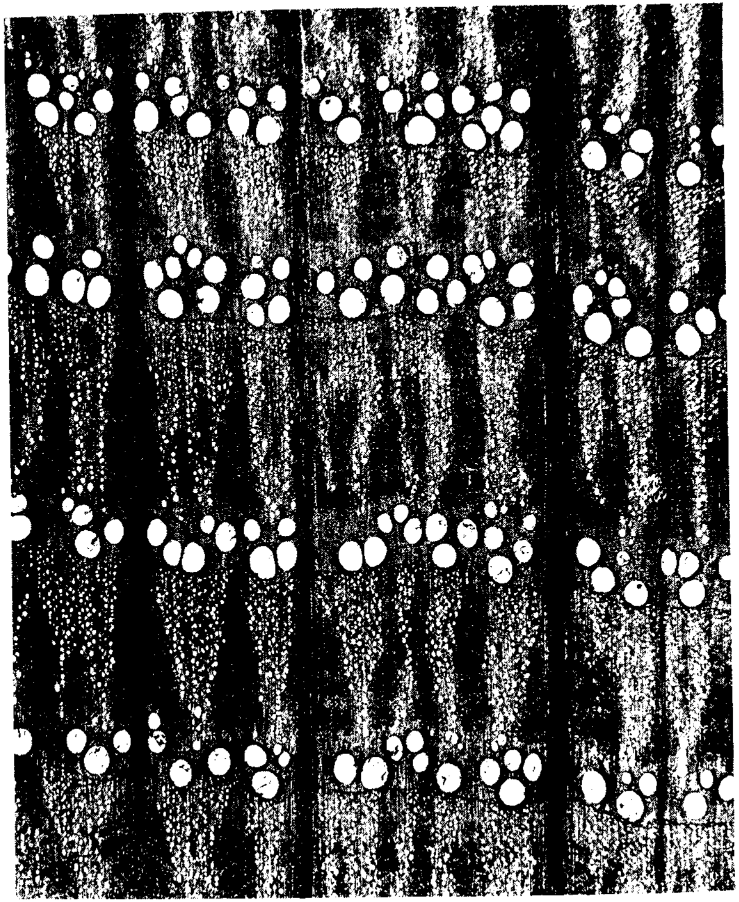
18–19

Ideas based on crystal structure were adapted for industrial design by the Festival Design Group, formed on the initiative of Mr. M. Hartland Thomas, of the Council of Industrial Design. An example is shown as Inset 20, immediately below. The following five pictures are photomicrographs of wood which may also suggest possibilities for design.

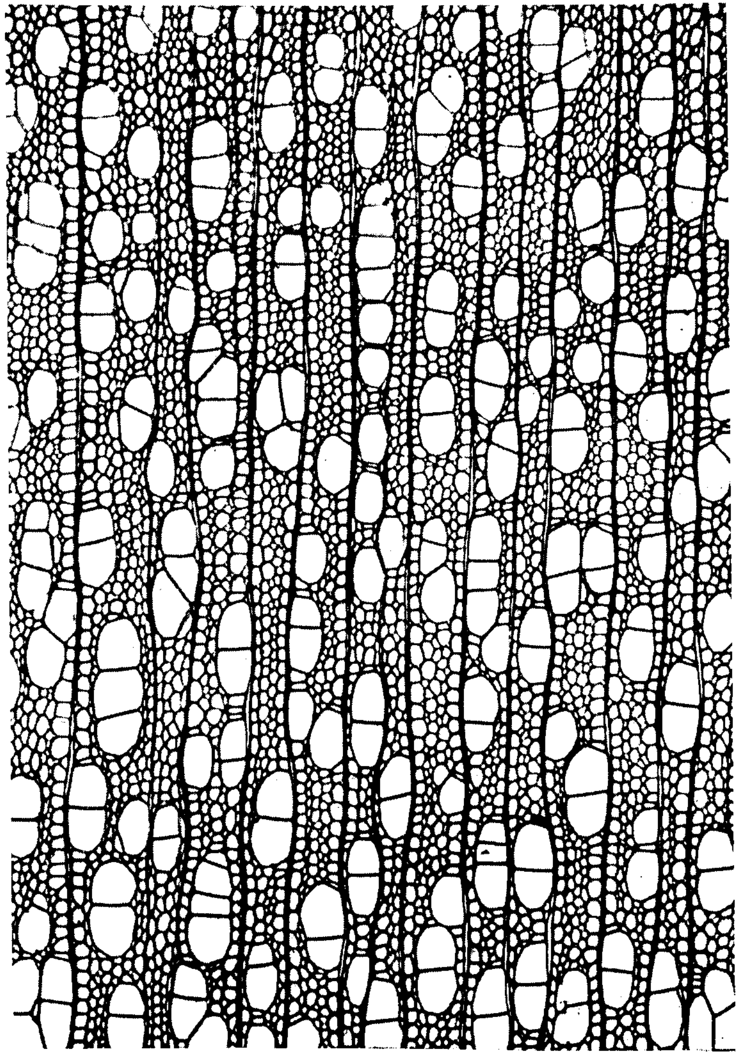
20. Sketch for warp tapestry, Warner & Sons, seen as heavy curtain in the Regatta Restaurant, South Bank Exhibition. *(By courtesy of the Council of Industrial Design.)*



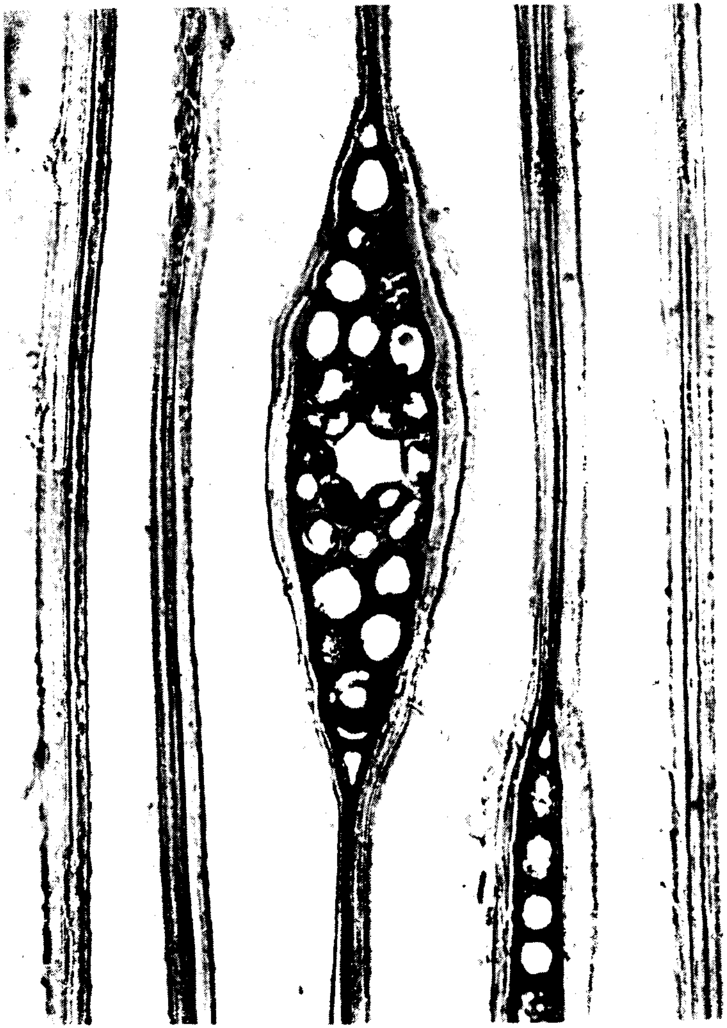
21. Cross-section of wych elm *(Ulmus glabra)*. Magnification × 11. See also following caption.



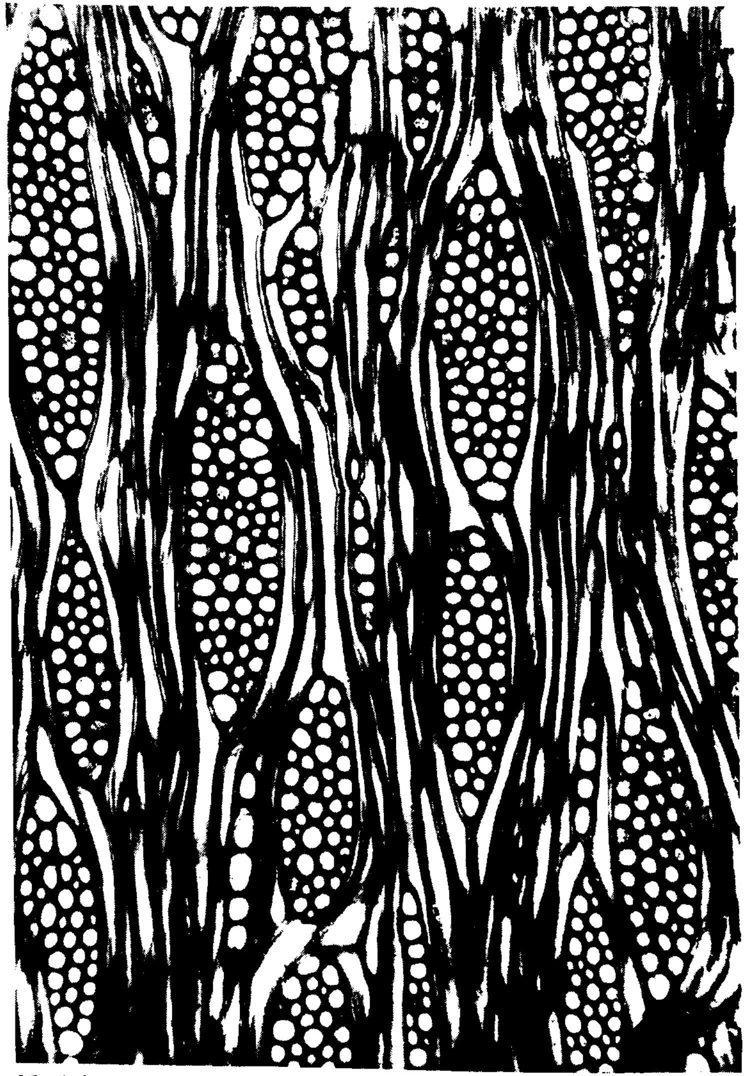
22. Cross-section of English oak *(Quercus robur)*. The white circles are sap-conducting pores. Lines of pores, cross-wise, mark the beginnings of annual growth-rings. There are numerous smaller pores in the later-formed portions. The narrow lines, from top to bottom in the picture, are rays; these radiate outwards and form part of the storage system of the tree. The smaller shaded areas are composed of fibres, which form the main structural framework. Magnification × 9. Inset 21, opposite, is so placed that the growth-rings run vertically in the picture. Wych elm is of similar structure to oak: among differences, the smaller pores are here grouped together to form a zig-zag pattern, and the rays are smaller, more numerous and less prominent.



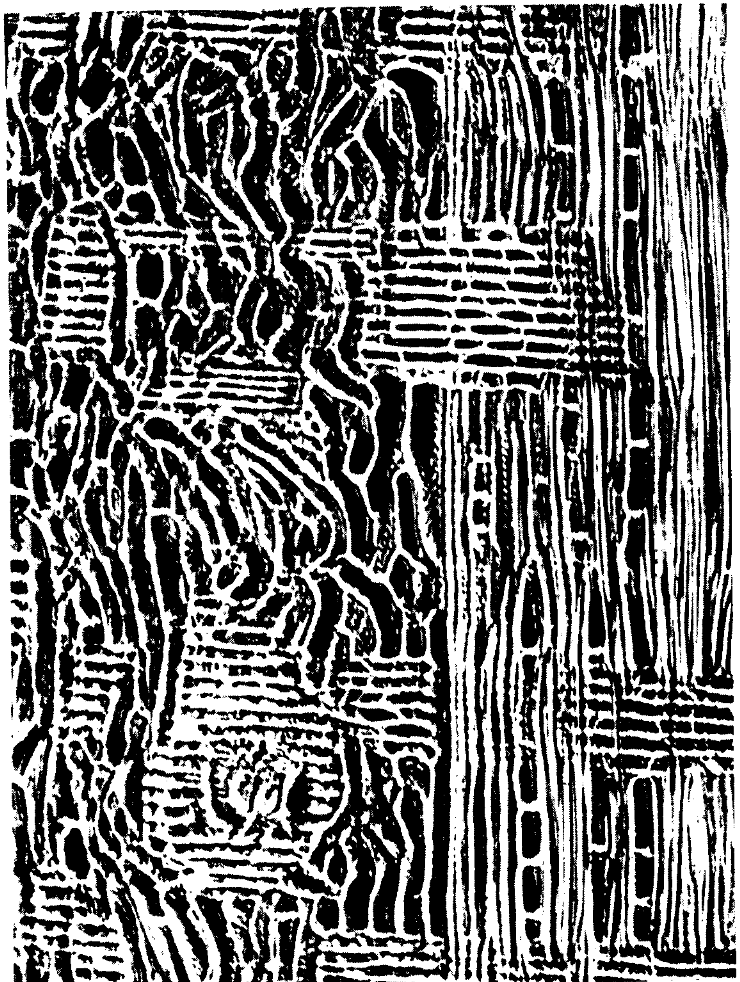
23. Cross-section of Canadian aspen *(Populus tremuloides)*, magnification × 23. An example of wood showing more uniform structure. Compared with Insets 21 and 22, pores are usually arranged in pairs or in bead-like strings, and are more uniform in size and distribution. Fibres are relatively thin-walled and do not show dark in the section.



24. Douglas fir, tangential section (i.e. vertical section parallel to growth-rings, and cutting across rays). The prominent boat-shaped structure of rays in the middle has been formed by the squeezing apart of the adjacent fibres. It is broader than usual, the large central hole being a horizontal resin-duct. The normal ray-structure is seen at bottom-right. Magnification × 400.



25. Ash *(Fraxinus excelsior)*, tangential section. There are again boat-shaped structures of rays, but without central resin ducts. The wide variation in the size of the rays, from 1 to 6 rows of cells in this case, is commonly encountered. Magnification × 150.



26. Radial section, English oak *(Quercus petraea)*. The tangled area in the left half of the picture represents a vertical section through the smaller pores of the later-formed part of a growth ring (*cf.* Inset 22). The upright lines, on the right, are fibres, and the brick-like structures part of the storage system. The groups of cross-lines are rays. Magnification × 130. (This and the preceding four pictures are published by permission of the Director, Forest Products Research Laboratory; *Crown Copyright* reserved.)

stance, uses only one tenth of the radio-frequency power of the 60-inch cyclotron. The lower dee-voltage also means some saving in the magnet, since large dee-voltages mean a large gap between the dee and tank walls. The much narrower gap which can be used with synchro-cyclotrons means that for the same amount of magnet-steel, and the same power consumption, a much greater magnetic field can be obtained. A change of design introduced as a result of the lower dee voltage and the necessity to modulate the frequency is that only one dee is used, together with an earthed sheath opposite its mouth, to concentrate the electric field. The use of only one dee simplifies the problem of modulating the frequency and has considerable mechanical advantages. It also leaves one half of the vacuum tank free for the introduction of auxiliary apparatus, such as the ion source, and for experimental apparatus and permits the placing of targets at varying radii, thus giving us different bombarding energies.

The frequency is modulated by means of a rapidly rotating variable condenser which is coupled into the dee with a resonant line. In the Berkeley 184-inch cyclotron the frequency is modulated from 11.5 megacycles per second at the moment of injection to 9.8 megacycles per second at the edge of the dee in accelerating deuterons to 200 million electron-volts or alpha-particles to 400 MeV. For protons of 350 MeV the frequency varies from 23 megacycles per second to 15.9 megacycles per second, and the protons travel some 100 miles during the acceleration.

In the conventional cyclotron particles are picked up from the ion source and accelerated continuously. In the synchro-cyclotron as the frequency is modulated the particles are captured into the stable orbit and accelerated only once during each modulation cycle. In contrast to the conventional cyclotron the beam from the frequency modulated cyclotron arrives in bursts, one burst for each modulation cycle. The modulation frequency is usually about 60 cycles per second, and the beam arrives in bursts of approximately 100 microsecond duration. The beam-intensity in synchro-cyclotrons is much reduced, but this is of little consequence at the very high energies produced. The arrival of the particles in short bursts does, however, complicate experiments with certain types of detecting equipment.

Several cyclotrons of this type are now in operation. The 184-inch cyclotron at Berkeley has been in operation since 1946. Four more are working in the United States at the Universities of Rochester, Harvard, Columbia, and Chicago, these machines producing protons of energies between 100 and 500 MeV. In Great Britain a 110-inch machine is working at the Atomic Energy Research Establishment, Harwell, giving protons of 170 MeV and a 156-inch machine is under construction at Liverpool (Inset 5–8).

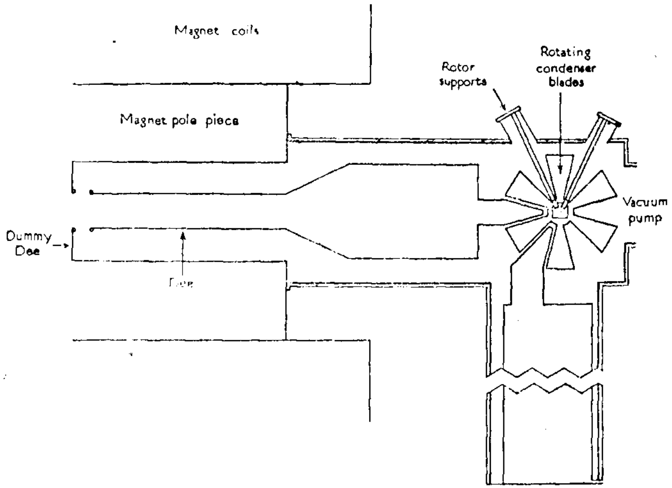


Fig. 4. University of Liverpool 156-inch synchrocyclotron: diagrammatic vertical section through the dee system. A photograph showing the dee of a 1/4 scale model is reproduced as Inset 7.

The radiations produced in synchro-cyclotrons are very penetrating, and radiation hazards present a considerable problem. Protons of 400 MeV have a range of some 13 inches in aluminium, and so can pass through the tank walls into the cyclotron room. This makes the cyclotron room dangerous to

personnel while the machine is in operation. The neutrons emitted under bombardment from these high energy particles are even more penetrating and are absorbed only by several feet of concrete. In order to bring the neutron-intensity down to a safe level in the area round the cyclotron the machine must be surrounded by many feet of concrete. In consequence synchro-cyclotrons are either surrounded by thick screens of concrete, or else are built completely or partially underground.

The 500 MeV protons produced in the largest synchro-cyclotrons are of considerably less energy than the energies of the most penetrating particles found in cosmic rays. The intensity of these very high energy particles in cosmic rays is very low, and for the study of fundamental particles in this very high energy region it is desirable to produce them in a laboratory. It is expected that experiments with particles in the BeV (1,000 MeV) range should provide results of considerable interest in connexion with the problem of the nature of nuclear forces. There is no theoretical reason why energies in the 1,000 MeV or BeV range should not be produced in a frequency modulated cyclotron. In practice there are very good reasons why synchro-cyclotrons of this type much larger than those already in existence will not be built. The magnet of the 184-inch cyclotron contains 4,000 tons of steel. To reach energies in the billion electron volt region, orbit diameters of 500 inches would be required and the cost of a magnet with pole pieces of this diameter would be prohibitively large and the power required for the continuous operation of the magnet would be very high.

PROTON SYNCHROTRONS

Fortunately, a more economical procedure is possible. The particles can be accelerated on an orbit of constant radius if the energy of the phase-stable orbit is increased by increasing the magnetic field. Using this method of increasing the energy permits the use of a ring-shaped magnetic field which results in a great saving in the total weight of the magnet. A proposed machine of this type to be built at Berkeley uses only 10,000 tons of steel as compared with the 4,000 tons used by the 184-inch cyclotron for producing particles of about 1/10th of the energy. Three machines to accelerate particles to the BeV region are under construction, two in the United States and one at Birmingham in Great Britain. These machines are called proton synchrotrons. The Birmingham machine has an orbit radius of 180 inches and is designed to produce protons of 1 BeV energy; the Brookhaven 'cosmotron' has a radius of 360 inches and is designed for 3 BeV protons and the Berkeley 'Bevatron' has a 600-inch radius and should produce 6 BeV protons.

In all machines the protons are accelerated on roughly circular orbits in a doughnut-shaped vacuum-box placed between the poles of the ring-shaped magnet. In the Berkeley and Brookhaven machines the magnet is built in four quadrants with straight sections between them into which equipment for

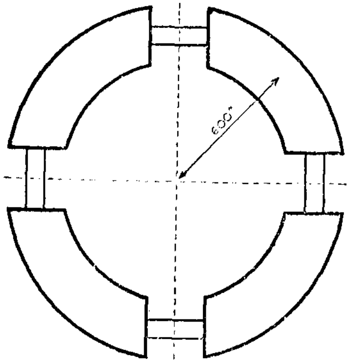


Fig. 5. The Berkeley proton synchrotron: simplified plan view showing the magnet-quadrants with straight sections between them (see text).

vacuum pumping and particle-injection is introduced. The design of the Berkeley magnet is shown diagrammatically in plan view in Fig. 5.

In the Birmingham machine the vacuum chamber is made of porcelain sections with a conducting layer of thin sheet-metal on the inside. The radio-frequency accelerating electrode, which

replaces the cyclotron dee, consists of a hollow metal conductor and covers some 95° of arc. In the Berkeley Bevatron the vacuum chamber is to be of stainless steel construction and the accelerating electrode is much shorter: in one of the straight sections the frequency range is from 0.34 megacycles per second to 9.6 megacycles per second and to maintain the phase stability conditions this frequency must at all times during the acceleration be accurately related to the increasing magnetic field. To achieve this large and accurate frequency change the acceleration has to be spread over a much longer time than in the synchro-cyclotron: it will take 1 second, compared with the 1/100th second of synchro-cyclotrons. Because the magnetic field is not constant, and it is therefore necessary to reduce losses from eddy-currents caused by the changing field, the magnet must be of a laminated construction and is built up from a large number of ½-inch sheets. This again means a long acceleration-time: for to vary the magnetic field from a few gauss up to 15,000 gauss in any period much shorter than a second would entail building the magnet of much thinner laminations. During this long acceleration the protons will travel a distance of some 200,000 miles, and during this time they must be confined in a chamber of only a few square feet cross-section. This means that the vacuum must be very good if the protons are not to be deflected from their paths by collisions with the few remaining air molecules in the chamber. This scattering by residual gas molecules is most important when the particles are at a low energy and so it is necessary to start the particles off in a proton-synchrotron with an energy of at least 1 MeV. Photographs of the Birmingham machine are reproduced as Inset 9–11. The Berkeley Bevatron (Inset 13) is to be injected with 10 MeV particles from a linear accelerator, while the Brookhaven machine uses a 4 MeV Van de Graaff electrostatic generator to inject the particles.

The particles are injected into the proton synchrotron when the magnetic field is low. They are held on the orbit by the radio-frequency field and slowly accelerated as the magnetic field is increased. After they have been injected the protons gain energy at the rate of about 1,000 electron-volts per turn. As they speed up, the magnetic field increases, and the time taken to make a circuit gradually decreases. To maintain the proper phase-stable orbit the frequency of the voltage applied to the electrode must be increased at the same rate. At first the frequency will be increased fairly rapidly, but as the protons gain speed the rate will be slowed down, and as the protons' velocity approaches that of light, the frequency becomes almost constant.

To increase the magnetic field from zero to its maximum of some 15,000 gauss every few seconds requires some 100,000 kilowatts. Fortunately the energy can be used over and over again instead of being drawn from the power supplies for each acceleration. The energy is stored in the large rotating flywheel of the direct-current generator which energizes the magnet. When the generator is energizing the magnet the flywheel slows down and the rotational energy of the flywheel is used in the form of magnetic energy. After acceleration, and before the next cycle commences, the energy flows back from the magnet to the flywheel to be used over again, so that only the frictional losses of all types involved in these transfers have to be supplied.

The output from these machines will consist then of bursts of particles of 1/10th microsecond duration and one pulse every 6 or so seconds. The intensity of the emergent beam will be small.

The construction of such machines is an engineering operation of large scale. The Brookhaven machine was started in 1947 and should be complete by late this year, the Birmingham machine was started a year earlier, and that at Berkeley in 1949.

Although none of these machines is yet in operation and there are engineering problems and uncertainties, there are grounds for considerable optimism about their future performance in that a quarter-scale model of the Berkeley Bevatron has been built and operated satisfactorily (Inset 12), so that all the principles of their operation have been tested.

PHOTOSYNTHESIS

H. LEES

THE Penguin *Dictionary of Biology* defines photosynthesis as 'The synthesis by green plants of carbohydrates from water and carbon dioxide with the aid of energy absorbed from sunlight'. This is a primary life process; the continuance of any life on earth depends on this unique ability of green plants, and of certain algae, to build up an organic compound from inorganic starting material without concomitantly breaking down some other compound in order to obtain the energy supply necessary for the synthesis. Any animal, for instance, can build up carbohydrates, but only if compounds provided in the food – including carbohydrates – are broken down to provide energy for the carbohydrate synthesis. Moreover the total synthetic process is inefficient. Only a fraction of the energy released by breakdown of the food is used for synthetic purposes: the rest is irrecoverably dissipated as mechanical energy or heat. In animals, and in most micro-organisms, the metabolic clock is running down; only in green plants and some algae is it wound up again. It is therefore not surprising that biochemists have made many attempts to elucidate the mechanism of photosynthesis; but unfortunately it has proved, on detailed investigation, to be so rapid a process involving compounds of so transient an existence that not until recently was any real understanding of it forthcoming. However, in the last few years, since the introduction of paper chromatography and the isotopic tracer technique, a very considerable body of facts has emerged. Of course, the picture is far from complete, but we can at least discern the outlines.

Any carbohydrate, such as glucose, cellulose, cane-sugar, or starch, can be represented by the formula $C_n(H_2O)_m$ where

m is equal, or very nearly equal, to n. For a hexose sugar (e.g. glucose) $n=m=6$, for cane sugar $n=12$ and $m=11$, for polysaccharides such as starch or cellulose, n and m may run up to hundreds. But if we put aside for the moment the question which of all the carbohydrates is the primary product of photosynthesis, we may write the general photosynthetic equation as:

$$n.CO_2 + m.H_2O + \text{light energy} = C_n(H_2O)_m + n.O_2 \quad \text{(Eqn 1)}$$

i.e. carbon dioxide, water, and energy are consumed, oxygen gas is released, and carbohydate is formed.

Early this century, F. F. Blackman was studying the rate at which photosynthesis proceeded in leaves under various conditions of light intensity, carbon dioxide concentration, temperature, and so on. He noticed that when the illumination was of low intensity, raising the temperature of the leaf had little effect on the photosynthetic rate; on the other hand when the illumination was intense the photosynthetic rate approximately doubled for every 10°C. rise in temperature. A doubling of the rate of reaction for every 10°C. rise in temperature is characteristic of normal chemical reactions involving the formation or disruption of molecules, whereas reactions independent of temperature are those involving some change in the structure of the atoms concerned. An example of a temperature-independent reaction is photoactivation; a substance absorbs light energy with the result that the electron-orbits at some point in the structure of the substance are excited to a higher energy-level; when the orbits return to their normal energy-level the energy they have absorbed is released again in some form; e.g. as chemical energy in a photochemical reaction. A familiar example of photoactivation is the exposure of a photographic emulsion. Subsequent development of the emulsion, however, involves only chemical reactions initiated by the release of chemical energy from the photoactivated constituents of the emulsion. In agreement with the temperature-sensitivities of the two types of reaction, the temperature of

exposure is irrelevant while the temperature of development is critical. On the basis of this difference in temperature-sensitivity, Blackman was able to interpret his results as indicating that photosynthesis involved at least two steps, a light-sensitive temperature-independent step, and a light-insensitive temperature-dependent step. According to this interpretation, a suitable formulation of photosynthesis would be:

Light-sensitive step ('light reaction')

$$\text{Photoactive compound } (z) + \text{light-energy} = \text{excited compound } (z^*) \quad \text{(Eqn 2)}$$

Light-insensitive step ('Blackman dark reaction')

$$z^* + n.CO_2 + m.H_2O = C_n(H_2O)_m + n.O_2 + z \quad \text{(Eqn 3)}$$

The two reactions form a sequence, in which the photoactive compound (z) provides the initial trap for light-energy, and the excited compound (z^*) provides the energy needed for synthesis. The z produced by the chemical reactions of Eqn 3 can, of course, be reactivated to z^* by a repetition of the light reaction, and the formation of carbohydrate by the reactions of Eqn 3 will go on as long as light-energy, carbon dioxide, and water are available. But just as the speed of the convoy is the speed of its slowest ship, so the speed of formation of carbohydrate in Eqn 3 will depend on the rate of the slower of the two reactions. If the illumination intensity is low, the rate of formation of z^* in Eqn 2 will be low, therefore increasing the *potential* rate of the chemical reactions of Eqn 3 by raising the temperature will not increase the *actual* rate since the rate of supply of z^* becomes a limiting factor. At high illumination intensities the rate of formation of z^* is high, and consequently no longer limits the actual rate of the chemical reaction; the rate of formation of carbohydrate now depends on the speed of the chemical reaction and is therefore temperature-sensitive. Which is what Blackman observed.

The correctness of this interpretation has since been confirmed by a rather different technique. If the green algae commonly used in photosynthetic experiments are illuminated steadily under suitable conditions of temperature, carbon dioxide concentration, and so on, the amount of photosynthesis that they carry out in a given time is proportional to the total amount of light that has fallen on them; in other words the total photosynthesis is proportional to the product of light-intensity and time of illumination. But if the light is now rapidly switched off and on repeatedly, all other conditions remaining unchanged, so that the total amount of light reaching the cells in a given time is now reduced, the rate of photosynthesis is not *correspondingly* reduced. The efficiency of the process, measured as the amount of photosynthesis produced by a given amount of light, increases. The explanation is that the photoactivation process is far more rapid than the subsequent chemical reactions. During a flash-period, when the light is on, the photoactivation process can gather sufficient energy (or accumulate sufficient intermediate compound z^*, which comes to the same thing) to keep the chemical dark reaction going for an appreciable period, and the chemical reaction will therefore be able to proceed whether the light is on or off. The rate of the chemical reaction will of course diminish when the light is off because the supplies of z^* are being depleted, but the reaction does not instantaneously stop, it continues for some time in the dark and therefore the amount of photosynthesis produced by a given amount of light increases. Under conditions of continuous illumination z^* simply builds up to a steady concentration and the rate of total photosynthesis is determined by the rate of the chemical reaction at this steady concentration of z^*. Light supplied in excess of the amount needed to produce this concentration is, if absorbed, uselessly dissipated as heat. These, and other similar, experiments on 'rate-limiting factors' leave no doubt that photosynthesis is at least a two-stage process involving two distinct types of reaction. We may now examine these reactions in more detail.

THE PHOTOCHEMICAL REACTION

Since only the visible band of the spectrum is appreciably active in photosynthesis the photochemical reaction must involve the absorption of visible light and must therefore be localized in the coloured part of the photosynthesizing organisms, in the green intracellular chloroplast bodies of green plants and algae. These bodies, which are fortunately separable from the rest of the cell and can be studied in isolation, owe their colour to the presence of chlorophyll, a green pigment with a chemical constitution not very far removed from the constitution of the red haemoglobin of mammalian blood (see *Science News,* **10,** 49). In addition the chloroplasts contain red carotenoid pigments. This is interesting from a biochemical standpoint since vitamin A, which is one half of a carotene molecule and is formed from carotene in the intestinal wall of animals, plays an essential part in vision (*Science News* **15,** 16); moreover the red light-sensitive eye-spot of certain unicellular flagellates contains carotenoids. But despite the obviously intimate connexion between carotenoids and photochemical processes in some living organisms, there are reasons for believing that in green plants and algae the photochemical reaction involves principally, if not solely, chlorophyll; since those wavelengths of light most strongly absorbed by chlorophyll are just those wavelengths that produce the most efficient photosynthesis in green plants and algae. In algae other than green algae it is possible that other pigments, including the carotenoids, are photosynthetically active. Perhaps indeed the carotenoids were the primaeval photoactive pigments which, in the course of evolution of green plants and algae, have been functionally although not physically replaced by the chemically more complex but photosynthetically more efficient chlorophyll. In animals, on the other hand, the carotenoids continue to play a photoactive role, while the photoactivity of the haem pigments (*Science News* **10,** 48) is completely subordinated to their oxygen-carrying properties such as we find developed in haemoglobin.

Just how the chlorophyll molecule is activated by light-absorption is not yet clear, but in some way the absorbed energy is used to split water into hydrogen and oxygen. The end result of this 'photolysis' of water is in some ways similar to the end result of the electrolysis of water in that both processes result in the release of oxygen gas. In photolysis the energy necessary to perform the split comes from light through the mediation of chlorophyll; in electrolysis the energy comes from an electric current. It is known that photolysis takes place in the chloroplasts because isolated chloroplasts evolve oxygen gas when they are illuminated. Moreover it can be shown that all oxygen evolved by the illuminated chloroplasts comes from water. If some of the molecules of the water in which the chloroplasts are suspended is 'labelled' with isotopic oxygen, i.e. if the water is made to contain an excess of H_2O^*, where O^* is an isotope of oxygen, then the evolved oxygen contains O^*_2 in precisely the same concentration at which H_2O^* is present in the water. Chloroplast photolysis differs however from electrolysis as normally observed in that hydrogen gas does not appear concomitantly with oxygen gas; in photolysis the hydrogen, instead of being released as gas, is used to reduce some compound X present in the chloroplasts. We may thus represent photolysis and electrolysis by the two equations:

Electrolysis

$$2H_2O + \text{electrical energy} = 2H_2 + O_2 \quad \text{(Eqn 4)}$$

Photolysis

$$2H_2O + 2X + \text{light energy via chlorophyll} = 2H_2X + O_2 \quad \text{(Eqn 5)}$$

This formulation suggests that photolysis by illuminated chloroplasts will go on as long as there is a supply of X, when all the X initially present has been reduced, photolysis should stop. This agrees with experimental observation; oxygen evolution from illuminated isolated chloroplasts soon stops but can be restarted if some compound capable of accepting hydrogen (i.e. some compound capable of replacing X) is added to the

chloroplast suspension. R. Hill, to whom most of our knowledge of photolysis is due, successfully used ferricyanide as a hydrogen acceptor; he found that ferricyanide added to an illuminated chloroplast suspension was reduced to ferrocyanide while a corresponding amount of oxygen was released. In practice Hill's experiments were not quite as simple as that, since it became necessary to prevent the photolytically released oxygen from reoxidizing the ferrocyanide to ferricyanide, but the technical details are irrelevant here.

Although we know that X can be replaced by ferricyanide, the nature of X remains unknown; moreover the situation is complicated by the fact that there is no reason to suppose that the X replaced by ferricyanide is the compound first reduced in photolysis. The reduction of X may involve any number of steps, e.g.:

$$2H_2O + 2A = 2H_2A + O_2 \quad \text{(Eqn 6)}$$
$$2H_2A + 2B = 2H_2B + 2A$$
$$2H_2B + 2C = 2H_2C + 2B$$

and so on where A, B, C, etc., are members of an oxidation-reduction chain ending in X

Such a formulation is not fanciful, similar hydrogen-carrier systems are commonly met with in biochemistry. Such chains are often linked with 'the cytochrome system' complexes of iron-containing haem pigments and protein (*Science News*, **10**, 68). It is therefore interesting that Hill has recently isolated a new cytochrome component, cytochrome F, from chloroplasts. Cytochrome F may well be an important component of the photolytic system, but so far its precise role remains obscure. However, by some carrier system, the hydrogen made available by photolysis is transferred to another centre of activity where it is used to reduce carbon dioxide. This reduction apparently takes place outside the chloroplasts, i.e. somewhere in the rest of the cell. Together with the hydrogen transfer mechanisms we have just considered, it constitutes the chemical or 'dark' reaction.

THE CHEMICAL REACTION

The simplest formulation of this reaction is:

$$2H_2X + CO_2 = H_2CO + H_2O + 2X \quad \text{(Eqn 7)}$$

where X is the last carrier in the hydrogen-transfer mechanism. The formulation has some historical interest because for a long time it was thought that H_2CO, otherwise formaldehyde, might be the first product of photosynthesis. This was because photosynthesis was at one time formulated as:

$$H_2O + CO_2 = H_2CO + O_2 \quad \text{(Eqn 8)}$$

E. C. C. Baly made many attempts to prepare formaldehyde, by the irradiation of carbonate solutions. He also tried to polymerize formaldehyde to hexose ($6H_2CO = C_6H_{12}O_6$) by similar methods. He may even indeed have succeeded, but we know now that, whatever success his experiments may have had, they could have shed little light on how photosynthesis actually proceeds, since there is little evidence that formaldehyde is produced at all and, if produced, it is certainly not polymerized to hexose under the influence of light. Carbon dioxide is not directly reduced to formaldehyde by the photolytically produced hydrogen; it is incorporated into a 'carboxylation cycle' and there reduced indirectly. The outline of this cycle, as far as it has been established, has been worked out principally by Calvin and his school in California. The technique they have used is almost childishly simple – the hallmark of a brilliant experimenter. Green algae were allowed to photosynthesize in the presence of carbon dioxide 'labelled' with a radioactive carbon isotope, i.e. the carbon dioxide contained some C^*O_2 where C^* is a radioactive isotope of carbon. After a suitable period of photosynthesis the algae were killed and their constituents, suitably fractionated, were separated on filter-paper as two-dimensional paper chromatogram (*Science News*, **13**, 23). This procedure separated the various compounds present in such a way that each compound could be

identified simply by the position it occupied on the filter-paper. The filter-paper, with the various compounds dotted all over it, was then placed over a piece of ordinary photographic bromide paper and left in the dark for a suitable period, after which the bromide paper was developed. During the period in which the filter-paper and the bromide paper were in contact, any compound containing radioactive carbon fogged the bromide paper beneath it by virtue of its radioactivity. Subsequent development of the bromide paper therefore revealed the presence and position of any such compound by showing a black blob. It was thus possible to identify, from the position of the black blobs on the developed bromide paper, the compounds into which the radioactive carbon of the carbon dioxide had entered during the photosynthetic peroid. The results of many hundreds of such experiments are now beginning to yield a fairly clear picture of what happens to the incorporated CO_2, and although different workers hold somewhat different opinions about the precise steps involved, the broad outlines are no longer in question.

The carboxylation cycle begins with a compound containing two carbon atoms. One molecule of CO_2 is added onto this to give a compound of three carbon atoms; this three-carbon-atom compound then accepts another molecule of CO_2 and is thus converted to a four-carbon-atom compound which then splits into two to give two molecules of the original two-carbon-atom compound and the cycle begins over again. The cycle does not, however, keep growing in volume at each revolution: it is tapped at the three-carbon stage and three-carbon compounds are diverted to other synthetic processes where, by a reversal of the paths by which carbohydrates are converted into three-carbon compounds (*Science News*, **10**, 61, and **19**, 116), they are converted to carbohydrate. The carbohydrates so formed are then used by the plant for the synthesis of proteins, fats, and other compounds it may require. The hydrogen from the photolytic reaction is necessary to keep this cycle going because the cycle is essentially a reductive one and hydrogen, as well as CO_2, must be continuously introduced into it. An example of a suggested carboxylation scheme is shown below:

$$\begin{tikzcd}
& \text{carbohydrate} \\
\underset{\textit{pyruvic acid}}{CH_3.CO.COOH} \arrow[ur] \arrow[r, "(+CO_2+2H)"] & \underset{\textit{malic acid}}{HOOC.CH_2.CHOH.COOH} \arrow[d, "(+H_2O)"] \\
\underset{\textit{acetaldehyde}}{CH_3CHO} \arrow[u, "(+CO_2)"] & \underset{\substack{\textit{glycollic acid} \\ \text{(two molecules)}}}{CH_2OH.COOH} \arrow[l, "(+4H-2H_2O)"]
\end{tikzcd}$$

The actual reactions are certainly more complex than the above scheme suggests. If acetaldehyde is one of the two-carbon compounds involved, then it probably occurs in a phosphorylated form as vinyl phosphate; phosphopyruvic acid and not pyruvic acid itself is the three-carbon compound; a one-step conversion of glycollic acid to acetaldehyde is unlikely, and so on. But despite these obvious objections there is little doubt that the scheme represents, in skeleton form, the sort of thing that happens.

Note added in proof. Since this article was written **W.** Vishniac and S. Ochoa (*Nature*, **167,** 768, 1951) have shown that an illuminated preparation of spinach chloroplasts can reduce carbon dioxide plus pyruvic acid to malic acid. This reaction, which uses di- or tri-phosphopyridine nucleotide as a hydrogen carrier (see *Science News*, **10,** 62), may thus be a key reaction in photosynthesis as suggested in the skeleton scheme.

FURTHER READING

Rabinowitch, *Photosynthesis.* Interscience Press, 1942. (This is difficult but repays study.)
'Symposium on photosynthesis' *Federation Proc.* **9** (2), 1950.
Stiles, W. 'Photosynthesis' *Science Progress,* **38,** 759 (1950).

THE ASSAM EARTHQUAKE OF 1950

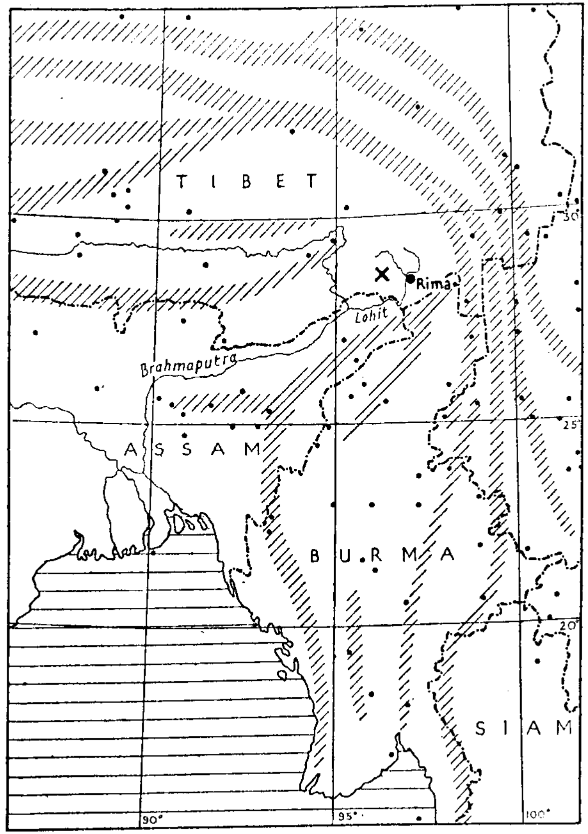
F. KINGDON-WARD

The Assam earthquake of 15 August 1950 was one of the two greatest, and possibly the greatest, since seismological observations became general. Mr Kingdon-Ward and his wife were in camp at the time among high mountains some 30 miles from the calculated epicentre of the earthquake. His account is reprinted from Nature *by permission of the Editors.*

ON 15 August 1950 my wife and I were camped in the village of Rima, on the left bank of the Lohit River, approximately lat. 28° 30′ N., long. 97° 00′ E., at an altitude of 5,000 ft. Rima is not really the name of a village, but rather of a small district containing five or six villages of rude timber houses and granaries, with a total population of a few hundred people.

Between the Rong Thö Chu-Zayul Chu confluence, a mile north of Rima, and the Indo-Tibetan frontier to the south, the Lohit gorge suddenly opens out, causing a sort of amphitheatre, some three miles from north to south, mainly on the east bank, where the villages are situated. The river itself is here more than 600 yards wide, twisting as though it were flowing across a flood plain. All around, the mountains rise very steeply from the flat terraces, a cluster of snow peaks on the Burma frontier, more than 19,000 ft high, lying within the twenty-five mile radius; but only four miles from the river they are already 15,000 ft high. Hence Western geographers might speak of the 'Rima basin'.

We had been several days in Rima, and had just concluded arrangements for sixteen porters to take us three marches up the La Ti torrent which, crashing down from the Burma frontier, flows into the Lohit immediately below Rima. We were to start early on the morning of 16 August for the alpine pastures.



Map showing the area of the Assam earthquake of 15 August 1950: **X** calculated epicentre; Rima, the immediate neighbourhood to which Capt. Kingdon-Ward's account refers; hatched lines, axes of mountain ranges; black dots, epicentres of previous earthquakes (after E. Tillotson).

It was about 8 p.m. by (our) local time, roughly three-quarters of an hour after dark on a hot, close night, the stars shining brightly up and down the arid gorge. My wife was in bed, our two servants were in their tent, and I was seated writing my diary near the entrance to our tent.

Suddenly, after the faintest tremor (felt by my wife but not by me) there came an appalling noise, and the earth began to shudder violently. I jumped up and looked out of the tent. I have a distinct recollection of seeing the outlines of the landscape, visible against the starry sky, blurred – every ridge and tree fuzzy – as though it were rapidly moving up and down; but fifteen or twenty seconds passed before I realized that an earthquake had started. My wife shouted: 'Earthquake!' before I did, and leapt out of bed. Together we rushed outside, I seizing the oil lantern which I placed on the ground; I was conscious of fearing that the tent would catch fire. We were immediately thrown to the ground; the lantern, too, was knocked over, and went out instantly.

I find it very difficult to recollect my emotions during the four or five minutes the shock lasted; but the first feeling of bewilderment – an incredulous astonishment that these solid-looking hills were in the grip of a force which shook them as a terrier shakes a rat – soon gave place to stark terror. Yet my wife and I, lying side by side on the sandbank, spoke quite calmly together, and to our two Sherpa boys who, having already been thrown down twice, were lying close to us.

The earthquake was now well under way, and it felt as though a powerful ram were hitting against the earth beneath us with the persistence of a kettledrum. I had exactly the sensation that a thin crust at the bottom of the basin, on which we lay, was breaking up like an ice floe, and that we were all going down together through an immense hole, into the interior of the earth. The din was terrible; but it was difficult to separate the noise made by the earthquake itself from the roar of the rock avalanches pouring down on all sides into the basin.

Gradually the crash of falling rocks became more distinct, the frightful hammer blows weakened, the vibration grew less, and presently we knew that the main shock was over. The end of the earthquake was, however, very clearly marked by a noise, or series of noises, which had nothing to do with falling

rocks. From high up in the sky to the north-west (as it seemed) came a quick succession of short, sharp explosions – five or six – clear and loud, each quite distinct, like 'ack-ack' shells bursting. It was the 'cease-fire'.

After about half an hour we returned to the tent. I noticed that the travelling clock on the table and my watch were both going; the altimeter still registered 5,000 ft exactly, and the thermometer outside under the fly showed 73° F. Nothing inside the tent was disturbed, except that a glass of water had been upset. Not a dog in the village barked. The tents were pitched on a small sandbank about 600 yards from the high bank of the La Ti torrent, rather more from the Lohit River. The nearest mountainside was 300–400 yards to the east of us, and, as it happened, that particular face did not slip badly; at any rate, no boulders reached us, or even the village. We could not have selected a safer site in an earthquake!

Within two hours the air was so thick with dust that every star was hidden; we breathed dust, it gritted our teeth, filled our eyes. Within a week every leaf of every tree and bush was caked with dust. Violent tremors continued all night, and twice we rushed outside – though the tent was perfectly safe. I slept intermittently for three hours; my wife not at all.

As soon as it was light, we dressed and went out to see what had happened. In the village every house had lost its roof of wooden slats, and every lean-to had been thrown down. The main irrigation channel was blocked, and the paddy fields with their standing rice would soon be dry. Farther afield, the small wooden monastery lay on its side, a bottle-shaped *chorten* had been stripped down to its axis, and a *mani* wall of flat inscribed stones had collapsed. I noticed that most of these structures had been thrown down towards the east, as though the push had come from the west.

Long fissures cut across the stony fields, running for the most part parallel with the river bank, past or present. In some places numerous fissures lay close together; elsewhere far apart. It appeared to depend on the nature of the soil, whether sandy or clayey. These cracks were rarely more than a few inches wide and two to four feet deep. Every well-trodden path was cracked throughout its length. Here and there a small block of land had sunk bodily.

The La Ti torrent was an extraordinary sight. The previous day the sparkling water, of a beautiful blue-green tint, revealed every boulder as it flowed swiftly down its rocky bed. Now it was dark coffee-coloured, but so thick with foam that it looked like *café-au-lait*. It had risen nearly four feet, and the muffled grinding of submerged boulders sounded ominous. The water smelt unpleasantly of mud, but I could detect nothing sulphurous about it. Three watermills, housed in timber cabins, had been smashed.

The Lohit had also risen, and the swift stream carried along great quantities of pine logs, which were already beginning to jam. The Rong Thö Chu was bringing down much more debris than the Zayul Chu, and the visible destruction up the former valley seemed to be greater than in the latter. (I have long held, contrary to the maps, that the Rong Thö Chu, which rises among glaciers, is the main stream. Its valley is also more thickly forested.) However, following up the Zayul River for half a mile, I soon reached a road block where a cliff had come down.

Day after day, night after night, tremors followed one another. Many of them were severe, and were always preceded by a short roar, like a distant thunder clap. The weather was sultry, the maximum shade temperature often more than 95°, until a wet spell between August 23 and 27 brought it down; the minimum was usually about 65°. On fine days a breeze blew from the south, but in the Rima basin it was so charged with fine dust, and so hot, that it brought no relief. By midday, as the rock avalanches began to fall, a vast dust cloud hid the sun, or dimmed it until it looked like a copper plate. Dust also mingled with the gathering thunderheads, and visibility was reduced to a mile or two.

No one was killed in Rima, but cattle and pigs were injured or killed by falling timber; in the river, great numbers of fish were killed, though few were taken from those raging waters.

All the hens stopped laying. The indifference, or fatalism, of the local population (nominally Buddhist) was superb. The day after the earthquake they went about their work in the fields. Yet many of them seemed to think there would be a repetition of the major shock, and that next time they might all be killed. We could find no one who recollected a previous earthquake. With their houses exposed to the weather, they slept in temporary shelters made out of boards; and mending their roofs in the evenings, our village at least looked normal again after a fortnight.

The rope-bridge across the Lohit had been broken, and until another rope was constructed and fixed in position we could not leave Rima. Our chief anxiety was potable water, and this proved something of a problem until we discovered a spring in the bank of the Lohit.

As for the mountains which enclose the basin, they had everywhere been badly mauled. Wide belts had been ripped off, carrying trees and rocks; whole cliffs had crashed down, deep wounds scored; and everywhere rocks continued to cascade down hundreds of gullies. The damage done in the main Lohit valley was bad enough; that done in the tributary valleys, where every stream had to break through a narrow gorge thousands of feet deep, was infinitely worse. The destruction extended to the very tops of the main ranges – 15,000-16,000 ft. above sea-level. No wonder the mountain torrents began to flow intermittently as the gorges became blocked, followed later by the breaking of the dam; whereupon a wall of water 20 ft high would roar down the gulley, carrying everything before it and leaving a trail of evil-smelling grey mud.

Everywhere the scraped cliffs glistened white in the sunshine. Fallen boulders showed that by far the most abundant rock is granite, containing tiny nuggets of black mica evenly distributed through it; occasionally the mica is in excess. Quartz veins ramify through the rock in all directions and are of frequent occurrence; and an abrupt change in the size and nature of the crystals, without any cleavage plane, is not infrequent. Not far from our camp was an outcrop of sandstone and schist with secondary white mica. A cliff of this had broken away and buried the main irrigation flume. But this type of rock was not met with elsewhere.

The Rima men made several attempts to fix the rope bridge and establish connexion with the right bank; but it was not until 6 September that they succeeded. Meanwhile no news came, and only a few wild rumours (based on occasional conversations shouted across the Lohit) reached us. When we left for Walong on 7 September, together with a patrol of the Assam Rifles who had joined us a fortnight after the earthquake, we had only a few days food in hand.

We reached Walong Outpost on 11 September after a risky journey, and any hope that below Rima the earthquake might have done negligible damage was quickly dispelled. (We did not yet know, of course, that we had been at the epicentre, or that we had survived one of the greatest earthquakes in history.)

It would be incautious at this stage to state categorically that the annual burning of the pine forest for the last fifty years, between Walong and Rima, was responsible for the huge damage done in the arid Lohit valley; but it seems certain that any further burning will prove completely disastrous. Even without the annual fires, the Lohit valley is likely to be a perilous place for some years.

HURRICANES

R. W. JAMES

TROPICAL depressions go by different names in different parts of the world. The Hurricane of the Caribbean area, the Typhoon of the China Seas, and the Baguio of the Philippines are all essentially the same type of disturbance. They are relatively rare occurrences, even in the regions most affected. Any particular locality is likely to escape a major visitation for years together, but when a settled area suffers the full fury of a hurricane, material damage may be reckoned in millions, and loss of life may be heavy. It has been estimated that 300,000 lives were lost in the Hoogly disaster of 7 October 1737; and, although disaster on this grandiose scale has never since been repeated, many later storms have exacted an appalling toll. It is therefore gratifying that during the present century the death-rate from hurricanes has declined dramatically in many parts of the world. In the area covered by the U.S. Weather Bureau Hurricane Forecasting Service, the mortality in severe hurricanes was one person to every $60,000 of storm damage during the five-year period, 1926–30; but during the five years 1941–5 the mortality was only one person per $2,500,000 of material damage. The ratio, loss of life to material damage, had been reduced in the proportion of 40 to 1.

The improvement may be attributed to better communications, superior hurricane-warning services, and increased co-operation between the authorities and the general public. The U.S. Weather Bureau has organized a hurricane reconnaissance service, with aircraft regularly patrolling at sea and reporting the characteristics and movements of disturbances otherwise inaccessible to direct observation. It would be ungracious, however, not to allow the hurricane specialist himself some share in the general improvement.

Weather-forecasting is still largely in the empirical stage. The value of a prediction is in some measure a function of the past experience of the man making it: for the behaviour of weather-systems is not known as a rigid sequence prescribed by initial conditions, but only in terms of statistically preferred patterns. There thus remains much scope for 'flair', or, to put it more boldly, for guesswork, inspired or otherwise. 'Flair' is nothing more than the subconsciously integrated past experience of the forecaster, and is most successful the longer that experience. It may be that in the future we shall know enough about the behaviour of hurricanes to cut out the guesswork in forecasting completely, but our present knowledge of hurricanes still falls far short of this objective. Painstaking investigation is more than ever necessary: the more we learn about the morphology, or form, of hurricanes, and their evolution, the better shall we be able to predict their effects on human life.

Tropical cyclones, as their name implies, achieve their most destructive effects in low latitudes. They virtually never occur closer to the equator than about 300 miles, and are usually born in latitudes 5–15° North and South, over the sea. In its early stages a tropical storm is of quite small extent, say 50–100 miles in diameter; but in that area wind-speeds are often more than 150 m.p.h. at the surface, and may be a good deal greater above the surface. The winds flow around the centre of the storm, where pressure is lowest, and spiral inwards. The moist sea-air which spirals inwards ascends from the surface, and is expelled from the vortex at a high level. It has been generally agreed that, as well as an ample supply of moist surface air, there must also be an unstable stratification in respect of temperature, for a hurricane to be formed.

In ascending, the moist surface-air expands and is cooled rapidly. Being cooler, it cannot retain all the moisture it first possessed. The excess moisture condenses out as cloud droplets, which may ultimately increase in size until they become large enough to fall to the surface as rain. In condensing, the excess moisture yields up its latent heat to the air. This latent

heat provides the source of energy which enables the hurricane to maintain itself, and to grow. Much of the kinetic energy of the wind is dissipated as heat by the frictional drag of the sea-surface on the air. This energy-loss must be made good from some source, such as released latent heat, or the wind-energy would soon be used up, and the hurricane would disappear.

The intensity of the hurricane will be regulated by the amount of latent heat released in the ascending motion. Hence we should expect intense rainfall to be associated with a developed hurricane. This is found to be the case, and, generally speaking, the more intense the rainfall the greater the development of the vortex.

It might be thought that the closer we get to the centre of a tropical vortex the more violent the display of the elements, but this is only true up to a point. At the centre of the hurricane there is a calm region, known as the eye. Here winds are relatively light, and variable in direction. The eye region in the early stages may be from 2–4 miles in diameter, about one-tenth the total diameter of the storm. The heavy overcast is missing from this region, the cloud-cover being thin, or even absent altogether. The inward-spiralling air does not reach the eye-region. However, the processes leading to the formation of a storm-eye are not fully understood. If winds were to increase at the rate they do in the outer vortex right up to the centre of the storm, the wind at the dead centre would have to blow at very great, probably infinite, speed from all points of the compass at once. This would be described by a mathematician as an essential singularity, and perhaps we might look upon the 'eye' as nature's way of avoiding so troublesome a concept.

The growth of a hurricane is normally arrested when it passes over land, and when cut off from sea-air the vortex usually dissipates. This is because a land-surface is rougher than a sea-surface, and the kinetic energy of the wind is dissipated at a greater rate. This loss could be compensated only by an additional energy-source, i.e. greater condensation in the rising air. However, when a vortex passes inland it tends to draw drier, continental air into its circulation, and so there is *less,* rather than more, latent heat released to make good frictional losses. Hence the hurricane weakens.

Hurricanes normally move at a speed of 10–20 m.p.h., in a northwesterly direction in the Northern Hemisphere. South of the equator the movement is towards the south-west. As time passes, a hurricane thus reaches higher latitudes. As it does so, the westerly motion gradually decreases, and the hurricane moves towards the east. This describes the normal track of a hurricane, but a forecaster can never rely on any particular disturbance to follow the normal path. A certain amount of work has been done on hurricane tracks, and it has been found that these disturbances are usually steered by the large-scale wind-flow of the earth, in particular by the circulation around the 'semi-permanent' oceanic anticyclones that inhabit the regions around latitude 30°. The 'recurvature' of the normal hurricane path is no more than the movement of the vortex from the equatorial side of the oceanic anticyclone, where the general wind-flow is from the east, to the pole-ward side, where the flow is from the west.

As the tropical depression reaches higher latitudes, its character changes. It expands outwards to cover a much wider area, and winds become rather less violent, although winds of hurricane intensity are still possible. The diameter of a mature disturbance in, say, latitude 35° may be of the order 500–800 miles, or about ten times the size of the immature hurricane of lower latitudes. At the same time the central region of relatively light winds expands to 50 or more miles in diameter, although this area begins to lose the sharp features of the eye in lower latitudes. Having reached the mature stage of development the disturbance may decay, or alternatively lose completely its character as a tropical disturbance, and become indistinguishable from a depression of the temperate regions. In many cases it has been possible to trace back a depression crossing the British Isles to its origin as a hurricane in the Caribbean area, but by the time it has reached Great Britain there is little left to suggest the furious vigour of its infancy.

The recognition of a general evolutionary development of

tropical disturbances is of great importance in forecasting their behaviour. Indeed, it might not be going too far to say that it is of comparable importance to the discovery by Norwegian meteorologists of an evolutionary process in temperate depressions. The 'polar front' theory, which for the first time emphasized the evolutionary aspects of temperate 'lows', has transformed forecasting in Europe, North America and elsewhere. A fuller recognition of the evolutionary development of hurricanes may lead to a similar advance in tropical forecasting.

If we are to understand the reasons for this particular type of evolution in hurricanes, we must consider the dynamic principles involved. We have seen that in a tropical cyclone, the air rotates about the centre, where pressure is lowest. The path followed by an individual parcel of air is thus curved, and to maintain that curved path, there must be a force on the air-parcel towards the centre of the cyclone: otherwise, it would travel in a straight line. The required force is provided by the excess pressure on the outer side of the air-parcel, compared with that on the inner side; much as the pressure exerted by the banked surface of a motor-racing track enables drivers to corner safely at speed.

Now let us suppose that we have a tropical vortex in equilibrium. Along each circle about the centre, there will be a succession of air-parcels moving at such speed that the excess-pressure provides the exact force needed to keep them moving in a circle of that particular radius. The excess-pressure will vary, with distance from the centre, in some specified way. Our job is to find out whether such a vortical motion is stable; meaning by this that any displacement of an air-parcel will lead to forces tending to restore it to its original position.

We can discover this by considering what happens to an air-parcel, supposed to have been arbitrarily displaced to some greater distance from the centre than its original position. The approach is to compare the speed of our air-parcel, after displacement, with that of existing air-parcels, which are already circulating at the same radius, and which we have assumed to be in equilibrium. If the speed of the displaced air-parcel is less than that of existing air-parcels at that radius, then the excess-pressure will be greater than is needed to keep it moving in a circle, and the air-parcel will tend to move back again towards its original position. On the other hand, if the speed of the displaced air-parcel is greater than this critical value, the excess-pressure will be insufficient, and the air-parcel will move further outwards: the vortex will be unstable. In practice, there is an effect caused by the rotation of the earth which tends to stabilize a vortex which otherwise would be unstable; but no essential principle will be sacrificed if we ignore this factor.

Let us, then, displace an air-parcel at point A in Fig. 1,

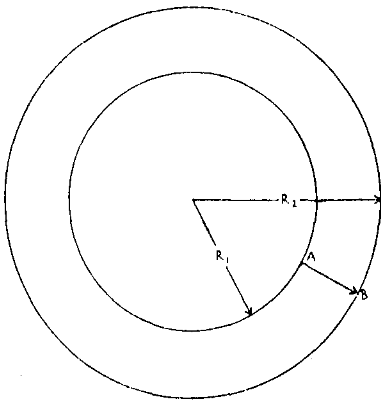


Fig. 1. To illustrate the problem of radial stability in a tropical depression. An air-parcel, moving on a circular path of radius R_1, is considered to be transferred from A to B, distant R_2 from the centre of the depression, and the motion of the air-parcel after transfer is then investigated.

moving at speed v_1 on a circle of radius R_1, to a point B at radius R_2. We may suppose that the only forces acting on the air-parcel are radial ones, so that its angular momentum, which is proportional to the product $V \times R$, will remain unaltered.

Hence, when the air-parcel arrives at B, it will have a speed v'_2, given by the equation:

$$v'_2 \times R_2 = v_1 \times R_1$$

so that

$$v'_2 = v_1 R_1 / R_2$$

This speed v'_2, it must be remembered, is that of the displaced air-parcel, not that of actual air-parcels at radius R_2. It falls off in inverse proportion to the radius, that is as $1/R$; and the condition for stability is that the speed of actual air-parcels should fall off more slowly than this proportion. The excess-pressure on the displaced air-parcel will then be greater than is needed to maintain it in a circular path at its new distance from the centre, and the air-parcel will move inwards again, as already stated. If wind-speeds fall off exactly as $1/R$, the vortex will be in *neutral equilibrium*: there will be no restoring force on a displaced air-parcel, but neither will it be unstable. On the other hand, if the wind-speed falls off more rapidly than $1/R$, the displaced air-parcel will have a greater speed than is consistent with circular motion in its new position, and the air-parcel will move to a still greater distance from the centre. We have an *unstable equilibrium*.

It is highly significant that in the early, immature stage of hurricane development the falling-off in wind-speed with distance from the centre is very rapid. At one point we have winds of hurricane intensity, but 25 miles away there may be a flat calm. Windspeed falls off more rapidily than $1/R$, and we have dynamic instability. In this situation, any slight outwards displacement of an air-parcel may lead, not merely to its further removal from the centre, but to its escape from the vortex altogether. In that way, the central pressure of the system may be progressively lowered, and the vortex intensified. In practice, the early, immature hurricane usually shows a very rapid fall in central pressure. There are thus strong grounds for supposing that the early growth-stage of a hurricane is connected with the emergence of dynamic instability. In a mature hurricane, on the other hand, although we may still encounter very high wind-speeds, the wind-speed falls off much more slowly with distance from the centre: the vortex is dynamically stable. Recognition of the initial instability may thus be a valuable tool in forecasting the later evolution of tropical disturbances through the working out of the instability.

The initial requirement of instability may throw some light also on the problem of why tropical cyclones do not occur more often than they do. The need for an ample supply of moist, oceanic, surface-air, combined with an unstable stratification in respect of temperature, has been mentioned earlier. These requirements are met very frequently over tropical seas, and it has been difficult to see why hurricanes should not occur more often. The additional requirement of dynamic instability is, however, more difficult to satisfy, and consequently the possibility of hurricane-development is more restricted.

Another feature of tropical storms, about which comparatively little has been known until recently, is the vertical structure of these disturbances. Thirty years ago it was believed that hurricanes were relatively shallow disturbances, the vortical circulation dying out above a few thousand feet. This view has been proved untenable for a number of reasons. It now appears that the wind-speed in hurricanes remains practically steady up to a height of about 10,000 ft, and thereafter falls off relatively slowly. In some cases the wind at 40,000 ft appears to be still half as great as that near the surface.

Although a hurricane in its early stage of development shows the greater rainfall intensity and higher winds, its total kinetic energy is small compared with that in a mature disturbance. This is because the high-wind area is so much more extensive in the latter phase.

The total kinetic energy developed by an immature disturbance may be of the order of 10^{23} ergs, while in a mature system we might have a total wind-energy of 10^{25} ergs, or say 100 times as great. For an intense depression of the temperate latitudes, a typical wind-energy would be 10^{26} ergs, about 1,000 times that of an immature hurricane. For comparison, it has been estimated that the total energy released by an atomic

bomb of the Hiroshima type was 10^{21} ergs. On this basis, the kinetic energy of an immature hurricane is about 100 times that of an atomic bomb, while an intense 'Low over Iceland' may contain the energy of 100,000 such bombs. The superiority of nature over man, in the matter of energy-release, is thus strikingly illustrated; and the difficulty suggested of any large-scale control of weather.

REFERENCES

Sawyer, J. S. *Q. J. Roy. Met. Soc.*, **73**, 101 (1947).
James, R. W. *J. Meteorology*, **8**, 17 (1951).

RESEARCH REPORT

A. W. HASLETT

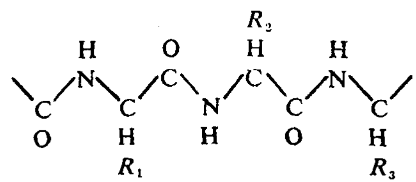
PROTEIN STRUCTURE

It is commonly agreed that the problem of protein structure can be solved by no one method, nor in a single step. A general discussion was given by J. C. Kendrew in *Science News* 11, and the present note is in the nature of a postscript on a single aspect: the configuration in space of the long chains, formed by the linking together of successive amino-acids, which seem to be a common feature of all proteins, fibrous and otherwise. As then stated, the chain 'might be straight: it might be coiled: it might be arranged in zigzags'. There was a multitude of 'hypotheses and speculations of varying degrees of plausibility', of which the two most widely accepted were shown in a diagram. There is now evidence which, for considerable groups of proteins, points to a form of coiled structure as the most likely. The evidence consists of a theoretical analysis by Professor Linus Pauling and others, of the California Institute of Technology, leading to a calculated model that meets the conditions from which they started; a favourable measure of agreement between the positions and intensities of *x*-ray reflexions from synthetic chains of similar type and those calculated from the model; and, finally, the confirmation by Dr M. F. Perutz, of the Cavendish Laboratory, Cambridge, of an additional *x*-ray reflexion that is predicted by this model, but by no other. This confirmation has been obtained both from one of the synthetic chains already mentioned and from a number of proteins. At the same time, it should be recognized that the model is not a universal one, nor put forward as such. In particular, ideas on the stretched configuration of wool or hair remain unaltered: the coil would be pulled out. Nor is chain-configuration the only aspect of protein-structure.

The existence of more than one configuration in proteins

was illustrated in a fundamental discovery by Professor W. T. Astbury: namely that different *x*-ray patterns are given by the fibres of hair, mainly keratin, according as these are stretched or unstretched when examined. The unstretched or α-configuration is characterised by *x*-ray reflexions corresponding with a repeat-distance of 5.1 angstrom units (5.1 A.) along the axis of the fibre. The same reflexions, and therefore repeat-distance, are found also in a number of other fibrous proteins, including those of wool, muscle, porcupine-quills, and bacterial flagella. They are found, too, in certain synthetic substances built up from amino-acids, and in partly synthetic fibres made from globular proteins, not initially fibrous; and indications of the same spacing have been found in haemoglobin, i.e. in an unaltered globular protein. It is a relatively small-scale spacing: 5.1 A. compared with the 57 A. diameter of the haemoglobin molecule, which is itself a relatively small molecule as proteins go. Aside from interpretation, there is therefore good reason for thinking that this α-configuration of unstretched keratin is quite a common configuration in proteins, though not unique. It is for such proteins that there is new evidence in favour of a coiled-chain type of structure; although it should be added for proportion that there have been other suggestions advanced by Professor Pauling on which the most recent evidence has proved unfavourable.

A chain formed by the linking together of amino-acids can be written diagrammatically as shown below, the main chain being represented by the zig-zag pattern of atoms stretching from left to right:



In this representation, R_1, R_2, R_3 etc. denote side-chains, derived from the original amino-acids, and irrelevant for the present purpose. Linkages of the kind CO–NH, as at the extreme left, are the peptide linkages formed by the joining up of amino-acids, the CO-group coming from the acidic part of one amino-acid, and the NH-group from the amino part of another. The remaining parts of the amino-acids, including the side-chains, are known as residues, since they are no longer amino-acids when joined together.

The distances between successive atoms of the main chain can be taken from the structures of individual amino-acids and simple derivatives which have been fully determined. Beginning with the CO–NH linkage at the left, these distances are respectively 1.32 A., 1.33 A., and 1.47 A.; and then the same sequence repeated. These figures, and those for the angles between bonds, which are similarly determined, are one necessary starting point; and in this respect the only difference between Professor Pauling's approach and that of others is that the figures have been more rigidly followed. As a further criterion he assumes that there must be additional linkages between each NH-group and some further CO-group which, though remote from it along the chain, could be above or below it in a coiled structure–as, for example, that of a coiled spring. This is by the comparatively weak type of linkage known as hydrogen-bonding. It is a condition which is met in part by various types of folding, but only in a continuous coil can it be met for all NH-groups. This was no doubt a preliminary consideration in deciding to examine the coil type of structure (formally a helix) in greater detail. Pauling and Corey found that there are five kinds of coil, and five only, in which the distances between atoms, and the angles between them are precisely as specified; and of these five there were two only which met this further criterion. In one of them, there are 3.7 amino-acid residues per completed turn of the coil, and in the other 5.1. There is little evidence to show whether the second of these coiled configurations is to be found in nature, though it meets the same conditions as does the first.

In the first type of coil, with 3.7 residues per turn, the 5.1 A. spacing of the α-structure appears naturally enough, and without further assumptions. It is the shortest or perpendicular distance between successive turns of the coil. In a crystalline material, it would be expected that the 'ideal' configuration would be so modified as to provide an exact repetition: that is, that there should be an exact number of amino-acid residues in every so-many exactly completed turns of the coil. The nearest whole-number coils are those with 11 residues in 3 turns, 15 in 4, and 18 in 5; and there are synthetic fibres, formed from amino-acids, of which the *x*-ray patterns are consistent with the first and third of these arrangements.

Thus far, it might be said that the evidence was plausible but not convincing. The new evidence, from Dr Perutz, relates to the repeat-spacing along the coil, the distance separating any point on one residue from the corresponding point on the next: for example, the C of one CO group from the C of the next CO group. In the 3.7 residues model, this distance is 1.5 A. It has been confirmed experimentally in *x*-ray reflexions from one of the synthetic fibres already mentioned, and also from horse-hair, muscle, porcupine quill-tip, and haemoglobin. It had been reported before, but not emphasized, in reflexions from porcupine quill-tip. The effect of the new evidence is to suggest that a 1.5-A. spacing, approximately across the fibre, is a regular accompaniment of a 5.1-A. spacing, approximately along it. There are, therefore, two tests instead of one for the α-configuration; and, from Dr Perutz's analysis, only a coiled structure can provide even a near approach to the 1.5-A. spacing, and only the 3.7-residues coil of Professor Pauling and his collaborators provides it exactly.

Among models previously proposed, the one giving the nearest approach was a coiled configuration with exactly 4 residues per completed turn. This would give a spacing of 1.4 A., instead of 1.5 A. It was one of the possible configurations considered by Sir Lawrence Bragg, J. C. Kendrew, the author of the *Science News* article, and Dr Perutz during a comparison of various suggested models, including some of their own, with the evidence from *x*-ray reflexions as [illegible] known. But the critical evidence was not then available to enable a decision to be made, although with further refinement this might have led to the 3.7-residues model.

There is one further point of interest. The structure of haemoglobin was earlier interpreted by Dr Perutz as corresponding with a squat cylinder, divided into four discs, and with the discs crossed by chains to which he attributed the α-configuration. However, the assignment of chain-configuration was necessarily tentative; and the confirmation of the 1.5-A. spacing in haemoglobin is taken as further evidence that it was correct. The case is thus strengthened for the recognition of an underlying similarity between the fibrous group of proteins that show the α-configuration, and the globular proteins which in gross structure could hardly be more different.

Pauling, L., Corey, R. B., and Branson, H. R. *Proc. U.S. Nat. Acad. Sc.* **37,** 205 (1951).

Pauling, L., and Corey, R. B. *Ibid.,* **37,** 241 (1951): with six accompanying papers by the same authors.

Perutz, M. F. *Nature,* **167,** 1053 (1951).

Huxley, H. E., and Perutz, M. F. *Nature,* **167,** 1054 (1951).

Bragg, W. L., Kendrew, J. C., and Perutz, M. F. *Proc. Roy. Soc.* '*A*', **203,** 321 (1950).

ELECTRIC FISHES

Research on electric fishes is of interest for a number of reasons. The most obvious is the high degree of specialization reached in particular species. An electric eel of 3–4 ft length can deliver a 600-volt shock through the discharge of some 6,000 batteries of plates, arranged in series down its length. About five-sixths of this length consists of electric organ. A torpedo-fish, with some 400 batteries of plates, gives a 40-volt shock. The thickness of an individual plate is about 50–100 micron in the electric eel, about 10–12 micron in the torpedo-fish. The degree of synchronization, most striking in the electric eel, is shown by the fact that, whereas the discharge of a single plate occupies about one thousandth of a second, the discharge of the whole 6,000 batteries disposed over several

feet length, is effected within about twice this time. This aspect of the matter, the close synchronization and the means by which it is effected, has been discussed lately by Professor A. Fessard of the Collège de France, Paris, during a series of lectures given at University College, London, as a visiting lecturer of London University. He has worked on the electric eel *Gymnotus* * in the Institute of Biophysics, Rio de Janeiro, of which the director is Professor C. Chagas; and earlier on the torpedo-fish *Torpedo marmorata* of southern Europe.

Second, there is a problem in evolution, which was noted by Darwin: namely that the production of a sufficient discharge to be of biological value seems to imply some considerable degree of organization before the stage of value was reached. There has been recent research by Dr. H. W. Lissmann at Cambridge which may bear on this problem. He has worked with species in which, instead of a controlled discharge, there is a continuous emission of weaker electrical pulses; the behaviour of the fish being modified either by changes in electrical conditions or by the emission of pulses similar to its own. This may have been the original function, although there is an obvious gap between this weak, automatic type of discharge and one under nervous control.

A third point, also mentioned by Darwin, is that between the electric organs of electric fish and 'ordinary muscle' there exists 'a close analogy in intimate structure, in the distribution of nerves and in the manner in which they are acted on by various reagents.' This analogy has been given greater precision. Common features are the use of glycogen as a source of energy, the release of acetylcholine at nerve endings, the presence of similar enzyme systems to those in muscles, and the attainment of a similar overall efficiency, about 20–30 per cent. The establishment of these similarities has been shared by a number of workers: among others by Dr W. Feldberg, now of the National Institute for Medical Research, London; Dr D. Nachmansohn, of New York; and, in relation to enzyme

* More correctly, *Electrophorus electricus*. '*Gymnotus*', though commonly used, is a legacy from an earlier classification.

systems, by Dr Couceiro of the Rio de Janeiro institute. There is also a similarity in electrical constants, for which Professor Fessard quoted figures. Electric capacity is of the order of 20 microfarad per sq. cm. in the plates of the electric eel, and about 5 microfarad per sq. cm. in the torpedo-fish; the corresponding figures for nerve membranes and muscle being 1–2 microfarad and about 10 microfarad respectively.

In his discussion of the synchronization problem, Professor Fessard pointed out that, on the simplest approach, it would be expected that discharge should begin near the head, and travel towards the tail at a rate governed by the extra distance. In the spinal cord of an eel, nerve impulses are transmitted at 40–50 metres per second. Even at the higher figure, there should thus be an interval of at least 20 milliseconds (ms.) between the stimulation of the front and the rear ends of the electric organ, compared with the 2 ms. duration found for the whole discharge.

Activity, as shown by experiment, follows a more complicated course. It begins commonly in a region nearer to the tail of the fish than to its head, and further removed, therefore, from the controlling higher centres; and from this starting point there is a wave of activity which travels in both directions. There are several contributory factors. One is that activity is helped, although not in itself produced, by currents originating in areas already active. Second, and connected with this, the units of the electric organ nearest to the head are those most resistant to stimulation. These areas, like the rest, need the primary stimulus of a nerve impulse. But because of their resistance to stimulation, currents directly transmitted from other areas may provide the final trigger-action. A second possibility is the use of varying delay-circuits, as would be the first thought of an electronic engineer if confronted with a similar problem. In fact, variations in the right direction have been shown in the laboratory, in response to artificial stimulation. Measurements have been made of the delays introduced at each of four synapses, or nerve-relays, which serve different regions of the electric organ from head to tail. Typical figures

are 3.5 ms., 2.8 ms., 2.0 ms., and 1.3 ms. respectively, so that delays are reduced as the distance to be covered becomes greater. Finally, there is help also from the thinning of the electric organ towards the tail, so that the time is reduced which is needed for lateral transmission across it. The two last factors, together, account for about 6 ms.; and, allowing for all factors, there remains only a balance, estimated at 1.6 ms., to be accounted for. This, Professor Fessard suggests, could be explained by small differences in the time of activation of the individual nerve cells in the higher centres, by which stimulation is initiated.

In the torpedo-fish, there is a less severe problem in synchronization. And at the opposite extreme there are species which could give a shock of the same order of magnitude, if synchronization had been evolved. This is not the case, and there is the reverse problem of explaining the relatively high degree of uncompleted, or possibly lost, specialization which is thus largely wasted. This brings us to the Cambridge experiments, which it has been suggested may throw light on the origin of the electric organ.

Dr Lissmann's first experiments were on an African freshwater fish *Gymnarchus niloticus,* found in the rivers of West Africa and in the Nile. A specimen of 28 cm. length was received from the Gold Coast through Dr K. R. S. Morris. When observed in an aquarium, it attracted attention through the possession of a swimming action which could be executed either backwards or forwards with equal ease; and second, through a marked ability to avoid obstacles when swimming backwards. Since the fish was known already to possess tissues in its hind end and tail similar to the shock-producing organs of 'electric fish', it was thought worth while to search deliberately for electrical signals. Pulses of almost complete regularity were found to be emitted at a frequency which was affected by temperature but not by the degree of activity or otherwise which was shown by the fish. Between 21° and 31.5°C., the frequency varied between 258 and 318 pulses per second, with the highest frequency at a temperature of 28°C. With the recording electrodes close to the fish, the greatest signal strength (about 30 millivolts) was with the electrodes aligned parallel to the fish and near its tail. But at a distance of up to 100 cm. in tap-water, signals could still be detected. Similar but distinguishable, signals were later found to be produced by two further species. One of these, *Gymnotus carapo*, is related to the electric eel. The other, *Mormyrops boulengeri*, is a member of a family with paired electric organs in the tail, though of lesser strength. There are differences, both in pulse-shape and frequency, between these three species.

Sensitivity to electrical changes in its surroundings was shown in the case of the original species *Gymnarchus niloticus*. This was done by taking a bent copper wire and dipping its two ends, in turn, into the water near the fish. It showed escape reactions which were not produced by the execution of similar movements with non-conducting material. On the other hand, if the fish's own electrical pulses were fed back into the water, but at a distance from the fish, then its response was to locate the electrodes and to attack them. There was also a marked response to the introduction into the water of artificially generated pulses comparable in intensity and frequency with those of the fish.

Lissmann, H. W., *Nature*, **167**, 201 (1951).

ANIMAL POPULATIONS

Theories of the growth and control of natural populations were reviewed by A. G. Thomson in *Science News* 20 in relation to the insects of stored products, and by J. J. de Gryse in *Science News* 21 in relation to forest insects. The subject is one which lends itself more readily to discussion in general, if mathematically expressed, terms than to explanation in terms of known and measurable factors. This is the more so when effects from migration are involved and there may be a mixture of populations. Both points are illustrated in a recent study of cycles in the fox population of the province of Quebec. It has been made by Dr L. Butler, Assistant Professor

of Genetics in the University of Toronto through analysis of the pelt returns for coloured fox kept by the Hudson's Bay Company from 1916 to 1948.

Among the facts regarded as already established are the existence of a 9–10-year cycle in the numbers of many Canadian animals living south of the timber-line, and of a 4-year cycle in the case of animals living north of it. This is shown in the fur returns for a number of species, including lynx, marten, mink, and muskrat; but, in spite of numerous theories, there has been no substantial evidence up to the present time to show the underlying cause of either cycle. The object of Dr Butler's analysis was to answer a more limited question: the nature of the transition from one form of cycle to the other in the case of an animal which is found on both sides of the timber-line. For this purpose, the records from some thirty collecting stations, covering a range of about 1,100 miles from north to south, were grouped into six regions, of which two were to the north of the timber-line and four to the south of it. Use was also made of coat-colour, classified in three groups; silver or black, cross, and red, to test the extent to which local populations are self-contained.

The results provide considerable evidence of migration, which has been predominantly in a northern direction up to the present time. In three out of the four southern regions there has been a consistent 9–10-year cycle during the whole period. But in the two northern regions there was a 9–10-year cycle, characteristic of southern regions, up to 1927; and from then onwards a 4-year cycle, though with signs of the 9-year (southern) cycle persisting on fuller analysis. At the same time there was a marked increase in peak trappings in each of the northern regions. The numbers of white fox, shown for comparison, maintained the (northern) 4-year cycle at all stages, and for these there has been no increase in population, apart from cyclical changes. The suggestion is that, up to 1927, coloured fox were being driven by recurrent population pressure into the two northern regions, but without significant local breeding.

From 1927 onwards, however, northern populations of coloured fox had become established, so that the predominant cycle was as for white fox, though with migration from the southern regions continuing as a secondary effect. The remaining region occupies an intermediate position, both in respect of latitude and population cycles. It is that centred on Fort George, in James Bay, the southward-stretching arm of Hudson Bay. Here the southern, or 9–10 year rhythm, was followed up to 1942; but in that year a subsidiary peak was shown corresponding with the northern rhythm. This is attributed to southwards migration, corresponding with increasing population-pressure farther north during peak years. Variations in the proportions of coat-types are analysed in detail and support the migration theory. For example, the proportion of northern 'silvers' has decreased from about 7 per cent to 2.5 per cent, while that of 'reds' has increased from 68 per cent to 76 per cent. The observed changes, taken region by region, are inconsistent with inter-breeding within a single population, but are consistent with the theory of locally isolated populations, in which isolation has been broken periodically by migration.

Butler, L. *Can. J. Zool.*, **29**, 24 (1951).

TREES AND FUNGI

The resistance to fungal attack shown by the wood of living trees can be approached as a problem in comparative biology, or through organic chemistry. The biological problem has been discussed lately by Professor E. S. McDonough and others of Marquette University, Wisconsin. They recall that there are common wood-destroying fungi which cause no serious trouble in living trees, even when these are already exposed by wounding. From studies on a number of species, they suggest that there is a general correlation between the amounts of anti-fungal substances which can be extracted from woods in the laboratory and the resistance to fungi shown by fallen and cut branches of the same species. Compared with the western or

'hardy' catalpa (*Catalpa speciosa*), the wood of the American elm (*Ulmus americana*), of which fallen branches decay quickly, contains less of inhibiting substances.

A more advanced stage has been reached in the study of specific substances extracted from the bark of poplars by boiling in water. This has been reported by Dr J. Grosjean, of the Institute of Agricultural Science at Wageningen, in Holland, and by H. L. Klöpping and G. J. M. van der Kerk, of the Institute of Organic Chemistry, the National Technical Organization, Utrecht. Dr Grosjean showed first that an inhibitor, extracted in the first place from the balsam poplar (*Populus candicans*), could be obtained also from a number of further species of poplar; that it could be concentrated in the laboratory; and that it could kill fungi, and not merely check their growth. He pointed out that, whereas it was not unusual for plant extracts to show fungicidal activity, poplars seemed unusual in that activity was higher in the bark than in the leaves. More detailed research at Utrecht has shown that there are a number of substances which contribute to the effect. Some, including salicylic acid, are of relatively low activity, but others are highly active. One, concentrated by successive extractions with boiling water, ether and boiling benzene, has been identified as benzyl gentisate, a derivative of benzyl alcohol and gentisic acid, so named after gentian, from which it was first obtained. This is a compound which has not before been described. It has been isolated as small white, needle-like crystals, and the identification confirmed by synthesis. It is present to the extent of 0.3 per cent in the fresh bark of balsam poplar. In laboratory cultures, fungal growth has been totally inhibited by concentrations of 100 parts per million and 250 parts per million according to the conditions of test, using a number of standard test organisms. A second substance, not yet fully identified, is more than five times as active.

McDonough, E. S., *et al. Nature*, **166**, 1034 (1950).
Grosjean, J. *Nature*, **165**, 853 (1950).
Klöpping, H. L., and van der Kerk, G. J. M. *Nature*, **167**, 996 (1951).

GULF STREAM SURVEY

A preliminary account has been given of a survey of the Gulf Stream, covering the 1,200 miles from Cape Hatteras, North Carolina, to the longitude of the Grand Banks of Newfoundland. It was carried out by six ships, under the general direction of Dr R. H. Fleming, Director of the Division of Oceanography, the U.S. Navy Hydrographic Office, during the period 6 June 1950 to 23 June 1950. The first object was to obtain a general picture, in amplification of earlier and more localized surveys. Second, it was hoped to obtain information on a number of specific questions; among others, rates of surface movement along the stream, the rate at which the stream changed position, and the formation of eddies resulting from large-scale meanders in its course.

During the general survey, the six ships were spaced about 150 miles apart, and zig-zagged into and out of the current, remaining continuously under way. Temperatures in the upper 200 metres were continuously recorded, positions were fixed by radio navigation, and current-strength measured by an electrical method. In the western half of the stream, the band of maximum current speed averaged 4–5 knots, with a falling off to the east, provisionally estimated at about 15 per cent south of the Grand Banks. The position of maximum current was usually on the north-west, or continental, side of the 'warm core.' Average temperatures in the upper 200 metres of the central region of the 'warm core' ranged from 68.3°F. to 75.5°F., compared with about 55°F. to 60°F. on the continental side. The falling off in core-temperature was not uniform, but was characterized by bodies of warmer water which 'showed up at intervals along the stream'. Temperatures on either side of the warm core were relatively consistent, so that the variations in core-temperature suggested a 'pulsing action in the current', possibly caused by short-term variations in the strength of the Trade Winds. It is these, by piling up water in the Caribbean, which are the primary cause of the north-easterly movement which is the beginning of the Gulf Stream. Preliminary observations off Cape Hatteras showed a shift in the position of the stream in this region at the rate of 4.5 miles a day towards the south-east; and results during the survey as a whole can be accounted for by 'lateral movements of not over 11 miles per day'. From occasional observations at greater depths, it is concluded that the most rapid rate of shift must be accompanied by vertical movements which extend to at least 2,000 metres depth.

The development of a major eddy in the stream was followed during a 10-day period. It originated in a cold-water arm which projected some 250 miles into the stream from the continental side; took nine days to break off, from the time of its discovery, when it was already well developed; and thereafter swung lengthwise along the direction of the stream. Available evidence is that major meanders, and therefore eddies, are not formed until well after the stream leaves the continental shelf, and probably not before 600 miles thereafter.

Only a preliminary account has yet been given. There is a large bulk of weather observations which have yet to be analysed, as well as many coloured air photographs and notes which were taken as part of the survey.

Fuglister, F. C., and Worthington, L. V. *Tellus*, **3**, 1 (1951).

MOUNT WILSON AND PALOMAR

Work at Mount Wilson and Palomar Observatories, California, has been reviewed by Dr I. S. Bowen, director of both observatories. His account covers the first eight months of regular observation with the 200-in. telescope on Palomar Mountain, as well as the beginning of full-time work on a systematic 'sky survey' with the 48-in. wide-aperture Schmidt camera, described by D. S. Evans in *Science News* 19.

In the programme of the 200-in. telescope first place has been given to the measurement of the distances of extragalactic nebulae, beginning with the most accurate method, which can be applied only to the nearest of the nebulae. This depends on the identification in the nebulae of stars of regularly varying brightness (Cepheid variables), for, from measurements of their apparent brightness and period of variation, their absolute brightness can be calculated. With the 100-in. telescope it was possible to identify variable stars of the wanted type in only a dozen or so nebulae – those within 1,000,000 light-years' distance – and on this sample there has depended the link with other and less accurate methods, used at greater distances. With the 200-in. telescope the Cepheid method should be applicable to some 30 nebulae, within 2,000,000 light-years' distance; and during the eight months from November 1949 there have been repeated exposures on six additional nebulae. The next method of measuring distances depends on the apparent brightness of the largest supergiant stars which can be distinguished in particular nebulae. Its reliability can now be checked further; and, as Dr W. Baade has pointed out, it should be usable with the 200-in. telescope up to 10,000,000 light-years. At greater distances use is made of the tendency of nebulae to be distributed in clusters, some containing hundreds of nebulae. The assumption in this case is that the brightest member of a cluster will be of similar brightness to that in other clusters, of which the distances have been determined. There is therefore a further interest, for the estimation of distances, in Dr Bowen's statement that, in the 10 per cent of 'available sky' so far covered in the sky-survey, 'more than 350 clusters and compact groups of nebulae have been discovered'. Many of these are near the limit of the plates at a distance of 300,000,000 light-years. As a preliminary interpretation from the survey Dr F. Zwicky is quoted for the statement that there are 'tremendous variations in the apparent brightness of nebulae, according to the direction of observation'.

According to a more recent statement by Dr M. L. Humason, a start has been made with the 200-in. telescope on measurements of the apparent speed of recession of extragalactic nebulae from the shift in wavelength towards the red of their spectral absorption lines. Compared with the 100-in. telescope, the limit of distance at which such measurements can be made has been extended from 250 million light-years to 360 million light-years, the increase in apparent speed of recession being from 25,000 miles per second to 38,000 miles per second. The relationship, that velocity of recession is proportional to distance, has thus been extended to the most distant nebulae to which the method can be applied. The significance of this and other types of measurement which can be made on nebulae was discussed by Professor G. C. McVittie in *Science News* 21: the requirement is for statistical information, the collection of which must take time.

Dr Bowen records also that, within our own nebulae, the Galaxy, new methods of observation are being applied to the so-called 'planetary nebulae'. These are expanded, and still expanding, shells of gas, generally with a star at their centre, which it has been suggested may have been thrown out during the supernova type of explosion. With the great light-gathering power of the 200-in. telescope, plates can be taken through deep filters, transmitting light of only a single sharply defined waveband; and these, Dr Bowen states, are bringing out features and details not otherwise observable.

Recent work at Mount Wilson, mentioned by Dr Bowen, includes spectrographic observations on the sun which raise a possible doubt as to the cycle of nuclear reactions involving carbon by which the energy output of the sun, in common with most stars, has been thought to be maintained. No trace has been found of spectral lines which should have been observable if the proportion of carbon-12 to carbon-13 were less than 40 to 1. The proportion on the earth is about 80 to 1, and the effect of the supposed nuclear cycle would be to give a lower proportion. Estimates, based on laboratory experiments, have ranged from 50 to 1 to possibly 5 to 1, assuming that the core and atmosphere of the star are well mixed. The Mount Wilson results are not, therefore, decisive. On the other hand, there is evidence from 'red giant' stars which suggests that variations in the proportion of carbon-12 to carbon-13 is a significant factor in relation to the past history of particular stars.

Finally, there is a point about the name 'Palomar' which may be worth explaining. The mountain on which the observatory is built is correctly described as 'Palomar Mountain'; and the observatory was first called the 'Palomar Mountain Observatory'. However, the analogy of Mount Wilson proved too strong, and 'Mount Palomar', which is incorrect, came into general use. For some time now, the names 'Palomar Observatory' or 'Mount Wilson and Palomar Observatories' have been the recognized descriptions; the latter being formally correct, since the two observatories are run as a single unit, both administratively and for the planning of their work.

Bowen, I. S. *Pub Astronom. Soc. Pacific,* **63,** 5 (1951).

SYNTHETIC QUARTZ CRYSTALS

Synthetic quartz crystals of up to about 100 g. weight were shown by C. S. Brown and others of the research laboratories of the General Electric Company, and Dr W. A. Wooster and Dr Nora Wooster, of the Brooklyn Crystallograpic Laboratory, Cambridge, at a Conversazione held earlier in the year by the Royal Society. Whereas it had been thought by many that synthetic crystals would be most successfully grown from molten material, it has been found that better crystals can be grown from solution. This is done at a pressure of 1,000 atmospheres. In the more successful of two methods, there is a continuous transfer of material from broken or defective quartz crystals, through solution, to the quartz plates used as nuclei for recrystallization. The crystals used as raw material are placed at the lower end of a thermal column at a temperature of about 400°C., recrystallization taking place at the upper end of the column at a temperature which is lower by about 40 degrees. The growth of a 100 g. crystal takes from 10 to 28 days. The crystals, when formed, are fully adequate for cutting as crystal plates for use in the control of radio-frequencies; and by electron diffraction, it has been shown that the best of the crystal faces thus formed are more nearly perfect than those seen in natural quartz crystals, while by *x*-ray methods they have been shown to be comparable in texture with the best specimens. Although natural quartz is an important

primary constituent of many igneous rocks, in which it must have crystallized from molten material, there is also much natural quartz which is of secondary origin. Many of the best crystals are found in cavities and fissures in silicate rocks, silica having passed slowly into solution through weathering, and quartz crystals deposited as a secondary product, as in the present process.

MICROSCOPE RESEARCH

The application of the scanning principle used in television to provide a new form of optical microscopy is being investigated by Professor J. Z. Young and F. Roberts in the Department of Anatomy of University College, London, with support from the Nuffield Foundation. Among the advantages which it is hoped to realize are the use of a projection screen some 3 ft square; control of contrast to suit the object viewed; and the use of visual displays for pictures formed essentially in ultra-violet or infra-red radiation. The principle is to place in front of the eyepiece of the microscope a cathode-ray tube which, as in its television use, is scanned continuously by a narrow beam of electrons that covers the whole tube, or whatever part of it may be required, in a sequence which is regularly repeated. Radiation from the moving spot on the screen of the cathode-ray tube is then passed through the microscope in reverse, so that a reduced image of the spot is formed on the specimen. The reduction is proportional to the magnification of the miscroscope, as normally used. The amount of light which is transmitted is determined by the density of the specimen at the particular point which is being viewed at the moment. Since only one point in the specimen is being viewed at a time, it is a limitation of the system that any movement in the specimen should be at a slow rate compared with that at which scanning is taking place. The light transmitted falls on a photomultiplier cell. And the output from this, in turn, is used to control the intensity of the electron beam in a second cathode-ray tube which scans in step with the first. This is of the small, high-brilliance type used to give projected pictures, and provides the final image.

Other suggested advantages arise from the greater sensitivity of the photocell, compared with the photographic plate, and from the ability to feed the electrical output which it supplies to further electronic devices.

In illustration of the first, it is pointed out that at a wavelength of 2650 A. in the ultra-violet, the photocell has an advantage of not less than 100 to 1, so that for a given brightness of picture, lower intensities of irradiation could be used. There should therefore be less difficulty in the avoidance of changes in biological material from the effects of ultra-violet irradiation during examination. In illustration of the second point, which it is thought may be the most important, it is suggested that 'particles could probably be counted and sized at the rate of a million per second'. A preliminary model has been built which 'suggests that these advantages can be realized', and further development is being undertaken.

Young, J. Z. and Roberts, F. *Nature,* **167,** 231 (1951).

CORRESPONDENCE

THE SPARK DISCHARGE

In a recent article on 'Lightning and the Spark Discharge' by Dr R. D. Connor in *Science News,* No. 19, the importance of elucidating the mechanism of spark breakdown in gases at atmospheric pressure was emphasized. The views expressed in that article concerning fundamental mechanisms have prompted us in this letter to draw attention to some recent results obtained in this laboratory and in America, which throw a new light on the problem of the spark discharge. These results show, contrary to the opinion expressed in Dr Connor's article, that the Townsend theory, involving primary and various secondary ionization mechanisms, in addition to accounting for breakdown at low pressures, is also of fundamental importance for higher pressures and larger gaps (e.g. 1 cm. gaps at atmospheric presssure).

In view of its importance it is first necessary to explain more fully what is meant by the Townsend theory of breakdown. This theory is based on experiments of the following kind. Imagine two parallel plane electrodes at distance d apart immersed in a gas at constant pressure p with a potential difference V applied: If d and V are varied so as always to keep the electric field X $(=V/d)$ constant, then the ionization current i (of the order of 10^{-10} amp.) at the anode resulting from a current i_0 (emitted photoelectrically from the cathode by an external source of radiation) is found, at low pressures, to be given by an equation of the form

$$i = i_0 e^{\alpha d}/1 - \frac{\omega}{\alpha} e^{\alpha d}) \quad \text{(Eqn 1)}$$

where α and ω are constant under the experimental conditions.

In air when d is small, $e^{\alpha d}.\omega/\alpha$ is small compared with 1, because although finite, ω/α is very small and for all practical purposes the current reduces to $i=i_0e^{\alpha d}$, i.e. the graphs obtained by plotting the logarithm of i against the distance d are straight lines. Then, if i_0 is known, and the currents i at small distances d are measured, we can obtain a value for α from these graphs. On the other hand, when d is large, the factor $(e^{\alpha d}.\omega/\alpha)$ can be appreciable compared with unity, and measurement clearly shows that the current i increases with d at a faster rate than exponentially, so that the graph of the logarithm of i against d now curves upwards. Since α is known from the part of the graph which is a straight line, we can now obtain a value for (ω/α). Thus the two constants in the current equation are known. The distance d can be increased to a limiting value d_s, when a spark occurs at the static sparking potential V_s $(=Xd_s)$. The Townsend theory interprets these results as follows: α is the coefficient of ionization by electrons (the primary process) and ω/α is the coefficient of secondary ionization which can involve processes such as ionization of gas molecules by positive ions, electron emission by positive ion bombardment of the cathode, or photoelectric emission.

Now returning to Eqn 1, it can be seen that there is a value of d, say d_B, for which the denominator of the current equation vanishes, i.e.

$$1-e^{\alpha d_B}.\omega/\alpha=0. \quad \text{(Eqn 2)}$$

Analytically this means that i approaches infinity, but this, of course, has no physical significance. Townsend, however, interpreted this condition as meaning that a current i in the gap can be self-maintained and independent of any external source of ionization; because, when $e^{\alpha d_B}.\omega/\alpha=1$, each electron at the cathode will regenerate itself and the current will continue to flow even if the external source of radiation producing the initial i_0 is switched off. Eqn 2 is the Townsend criterion for breakdown, since in this condition the gas is in an unstable state and the current can proceed to a spark, glow, arc or other type of discharge depending on the external electrical circuit. Insertion of values of α and ω (obtained from pre-breakdown current measurements) in Eqn 2 gives a predicted value for the sparking distance d_B and therefore for the sparking potential $V_B=X.d_B$. These predicted values agree well with those observed experimentally at low pressures, i.e. $d_B=d_s$.

However, the application of this theory to breakdown at high pressures has met with three main objections, viz:

(i) No experiment had shown an upcurving of the graphs of log i against d; i.e. no evidence was found for the existence of a secondary ionization process building up the pre-breakdown current i.

(ii) Under certain conditions of applied voltage (about 2 per cent in excess of static spark voltage), breakdown occurs in times (about 1μsec.) so short as to preclude the movement of most of the positive ions across the gap to the cathode.

(iii) There was no dependence of sparking potential on cathode material: this was taken to indicate that no cathode secondary ionization mechanism existed.

These objections, which led in some quarters to the rejection of the Townsend theory and to the development of a fundamentally different view based on a streamer theory (outlined in Dr Connor's article), now have to be reassessed in the light of modern work.

In the first place, experiments[1] have shown that by measuring the current i sufficiently close to the sparking distance, a definite upward curving of the graphs log i against d has been revealed in air at the higher pressures. Thus objection (i) is no longer valid. The reason such curvature had not been observed before was that it had not been fully realized that, because ω/α is very small, the factor $e^{\alpha d}.\omega/\alpha$ only becomes appreciable when d is large; thus, to detect upcurving, measurements must be made close to the sparking distance where the experimental difficulties are very great. Values of α and ω obtained from the graphs, when inserted in Eqn 2, predict values of the sparking potential which are found to agree closely with those measured.

Secondly, recent measurements[2] have shown that, at voltages very slightly greater (about 0.05 per cent greater) than the sparking potential, the formative time lag is sufficiently long (about 100 μsec.) to allow positive ions to cross the gap and produce a secondary ionization at the cathode. The time lags decrease rapidly with increasing voltage, but it can be shown that a secondary process involving positive ion action and/or photo-electric effect at the cathode can account for such a rapid decrease. This result, therefore, disposes of objection (ii).

With reference to the dependence of sparking potential V_s on cathode material, the Townsend sparking criterion shows simply that, when ω/α is a function of X/p, cathode influence on V_s is extremely small at high pressures and no measurements with sufficient accuracy to detect any dependence have been published. On the other hand, when the cathode process is not a function of X/p, say a function of the field X only, theory indicates that strong dependence of V_s on the cathode is possible. Such a dependence has been observed as a marked departure from Paschen's law at very high pressures and short gaps, when the field X is high[3]: the high field, besides producing cold emission, can enhance the secondary emission due to the impact of positive ions.[4]

Consider now an alternative view, the streamer theory, which introduces a sudden process to account for spark breakdown at high pressures. This theory does not *quantitatively* explain the observed increase of current with distance given by Eqn 1. Furthermore, the streamer view is based on rapid processes which do not account for the observed long formative lags at the static breakdown potential. Finally, the streamer theory is based entirely on two gas processes, electron collision and photo-ionization, not involving the cathode.

This brief survey of the available evidence shows that the very points on which the Townsend theory was rejected now actually support it. In fact, the Townsend theory is the only one which gives a consistent picture of the development of the spark discharge from the growth of the small pre-breakdown ionization currents right up to the static breakdown condition. The relative importance of the several secondary mechanisms possible on such a theory has yet to be ascertained, and in view of the importance of this subject of high

pressure breakdown, both fundamentally and in its application to high pressure insulation for electrical engineering and high voltage nuclear physics, research into the remaining problems is being actively pursued.

J. DUTTON
S. C. HAYDON

Department of Physics, University College of Swansea

1. Jones, F. L., and Parker, A. B. *Nature*, **165**, 960 (1950).
2. Fisher, L. H., and Bederson, B. *Phys. Rev.*, **81**, 109 (1951).
3. Trump, J. G., Cloud, R. W., Mann, J. G. and Hanson, E. P. *Elect. Engin.*, **69**, 961 (1950).
4. Jones, F. Llewellyn, and Morgan, C. G. *Phys. Rev.*, **82**, 970 (1951).

Dr R. D. Connor writes:

The letter above is of importance as it has described clearly the main objections to the Townsend point of view, and also indicated some of the most recent work which can be interpreted as supporting that view, but before commenting thereon, I must discuss the earlier part of this letter.

Eqn 2 above is the same as my Eqn 2 in my article[1] with ω instead of β. ω is here taken as a generalized Townsend secondary coefficient which includes among other mechanisms 'ionization of gas molecules by positive ions'. This mechanism, I stated, did not occur in the gas for electric fields such as were present in the spark discharge. I have found no evidence for changing this view. Such experiments as have been performed to measure the threshold energy for positive ionization give values of the order of 100 volts and upwards for alkali ions in the rare gases.[2] No ionization by H_2 positive ions in argon is observed for energies up to 1,000 volts, and protons produce no evidence of ionization in hydrogen, helium or argon up to 4,000 volts. It would appear that alpha particles do not ionize air appreciably at energies several times this latter figure.[3] There is no known mechanism whereby a positive ion in a normal spark discharge can achieve anything like these energies.

Thus, unless some recent work overturning these earlier measurements has escaped the present writer's attention, the situation on this point remains as in *Science News* 19. A similar view has been expressed recently elsewhere.[4]

However, in the year which has elapsed since the writing of the article, work has been published which would appear to be interpretable in two different ways, firstly, as stated opposite, that Townsend mechanisms and theory alone are competent to explain all the observed phenomena, and secondly that secondary mechanisms while being active require the additional action of a streamer to explain the observed discharge. This latter view which is supported by Fisher and Loeb is the one which the present writer feels more inclined to accept. The evidence is as follows.

In the search for the transition from Townsend to streamer mechanisms no discontinuity was observed. Meek found that for values of *p.d.* of 200 mm. of mercury and less, and 1 cm. gap, streamer theory gave sparking potentials in excess of those observed and so suggested this as the transition region from streamer to Townsend mechanism. Fisher[5] has pointed out that by changing the variable in the calculation, satisfactory agreement on the streamer theory can be obtained well below a *p.d.* value of 200. Recently the Swansea group[6] have obtained evidence of the presence of Townsend mechanisms for *p.d.* values above 200. Even more recent work in the United States on the formative times of sparks in air has been published[7], in which it is shown that the formative time depends much more on the overvoltage than on the *p.d.* value. In fact, formative times are found to be independent of the pressure, and to be very long at very small overvoltages. The times decrease very rapidly with increasing overvoltage so that at 2 per cent overvoltage the formative time is of the order of 1 microsecond. For larger overvoltages, the breakdown in air is known to occur in very much shorter times.

The Townsend view is not shared by the authors of this work[5, 7] nor by Loeb.[8]

The authors suggest that the relatively long formative times

for overvoltages up to 2 per cent can be explained by considering the gradual build-up of space-charge in the gap by succeeding avalanches, engendered by electrons liberated from the cathode or gas by photons from the initial avalanche. The presence of the space charge increases the applied field locally thus increasing α (for α/p is an increasing function of x/p), which causes an effect similar to the action of a secondary mechanism. *The space charges increase until the charge-densities necessary for the formation of a streamer occur.*

Thus a modification of the existing streamer-theory is necessary as outlined above to account for these long formative times at small overvoltages. The modification has yet to be examined quantitatively. Fisher and Bederson[7] go on to point out that a space-charge mechanism can explain the long time lags observed at low and high pressures. They state that as the time lags at lower pressures for overvoltages of a few per cent are as short as those obtaining for high pressures, 'it is clear that the classical Townsend theory, hitherto assumed valid for *p.d.* less than several hundred is inadequate'. (It should be noted here that space charges as such do not enter as mechanisms in the Townsend theory.)

Loeb[8] has recently commented on this work and agrees that where formative times are long, the unmodified streamer theory is not applicable. The opinions in references 5 and 7 are here endorsed. Further, Loeb points out that on this view a unified theory of sparking can be built, with the threshold for space charge formation (ultimately leading to streamer type breakdown) being set by the Townsend criterion.

The reality of the streamer is not questioned as yet by any save the Swansea group. In fact, in a recent review Craggs and Meek[9] have outlined recent work in many centres on streamer-breakdown.

Much of the earlier theoretical work used comparison of calculated and observed sparking potentials as evidence for the correctness of the underlying mechanisms. It is now known that almost any reasonable criterion for breakdown can be made to fit the observations; hence no vindication of a mechanism should rest on sparking potentials alone.

It is pointed out in the letter above that streamer-theory does not account for the observed increase of current with gap length. It was never intended to do so. The streamer mechanism is operative only during the transient discharge. There can be little doubt that pre-breakdown currents, where these are measurable, are describable in terms of Townsend mechanisms with, perhaps, effects such as are mentioned above. It would be interesting to measure more pre-breakdown currents in long gaps at high pressure.

There are many aspects of the spark discharge which are difficult to reconcile with the Townsend theory taken by itself. Some of these are listed below.

(1) It is difficult to see how breakdown occurs in very long gaps and in the lightning discharge.
(2) Mid-gap streamers, such as are observed in cloud chamber photographs, corona-streamers from positive points, and the zig-zag channel of the long spark are difficult to visualize also.
(3) Positive and negative leaders are similar.
(4) The action of certain types of spark counters[10] for alpha-particles is much easier to visualize as a space-charge-streamer effect.

Thus, while not even the founders of the streamer theory would claim its finality, and while it appears established that Townsend mechanisms are active in regions where they were formerly not thought applicable, the streamer theory cannot be considered as overthrown. It may be that in the future the streamer theory will be entirely replaced by a new view of the events. In the opinion of the writer and others, that day has yet to come.

1. Connor, R. D. *Science News* **19,** 91 (1951).
2. Landolt-Bornstein (Tables) Vol. I p. 355 *et seq.* Berlin, Springer-Verlag, 1950.
3. Loeb, L. B. *Fundamental Processes of Electrical Discharge in Gases,* p. 375, John Wiley, 1939.

4. Wilkinson, D. H. *Ionization Chambers and Counters,* p. 49. Cambridge University Press, 1950.
5. Fisher, L. H. *Electrical Engineering,* p. 613. July, 1950.
6. Llewellyn-Jones, F. and Parker, A. B. *Nature,* **165,** 960 (1950)
7. Fisher, L. H., and Bederson, B. *Phys. Rev.,* **81,** 109 (1951).
8. Loeb, L. B. *Phys. Rev.,* **81,** 387 (1951).
9. Craggs, J. D., and Meek, J. M. *Research* **4,** 4. 1951.
10. Connor, R. D. *Proc. Phys. Soc. 'B',* **64,** 30 (1951).

CORAL REEFS

IN *Science News,* **8,** 116, it is stated that 'Darwin ... thought a fringing reef of coral would be established round the edge of the island – although, in fact, there are no known examples.'

Rotuma (12° S., 177° E.) is, however, such an island. It is hilly (rising to 840 ft), 8 miles by 2½ miles, has a fringing reef about 300–400 yards in width, and no barrier reef. The 100-fathom contour follows the outline of the island one mile or so offshore, except to the north-west and west where it runs out to some 5 miles to enclose three small islets (see Admiralty chart).

J. Stanley Gardiner, who was with the Royal Society coral reef expedition to Funafuti, and later occupied the chair of zoology at Cambridge, went on from Funafuti to Rotuma in September 1896 (in H.M.S. *Penguin,* be it noted). He reported on the island in 1898 (*Q.J. Geol. Soc.,* **54,** 1). He quotes Darwin as saying that 'it possesses a fringing reef only', and states: 'On the reef, which is fringing, *with in places an approach to the barrier class,* are several islets with a gradual slope towards the land, but precipitous cliffs to the sea.'

The words I have underlined rather confuse the issue and have no clear meaning unless he refers to the reef at its greatest width at the east end, where there are stretches of water remaining at low tide, i.e. the reef is not all dry or awash.

H. S. EVANS

Rotuma, Fiji

BOOKS RECEIVED

SCIENCE ITS METHOD AND ITS PHILOSOPHY by G. Burniston Brown (Allen & Unwin, 1950), pp. 189, 15s.

Dr Burniston Brown has approached the scientific method through the learning process in animals, a discussion of words and their use, and a historical sequence represented by Aristotle, Francis Bacon, Newton, and finally modern science, of which Whewell of Trinity, Cambridge, is taken as an early and articulate proponent. The definition is offered that '*Facts* are propositions which are *true*, and propositions are known to be true if they can be verified', the latter process consisting in 'the arrangement of situations in which there are associated events of consciousness'. The position of Eddington is considered sympathetically, but as an avowedly dangerous excursion. Dr Burniston Brown's great service is to convey at a relatively easy level that a subject which is often rendered in platitudes is full of difficulty.

A SIGNPOST TO MATHEMATICS by A. H. Read (Thrift Books, 1951), pp. 123, 1s.

This is an able and lively presentation, in extraordinately few words, of the essence of mathematics. Conceptions such as form, symmetry, rigour and generalization should be given some real measure of meaning to those who may have been uncertain in their use, and to the half-mathematician there may be a pleasure of reminiscence of the same kind as was given by E. V. Lucas at his best to the half-traveller.

THE EARTH'S CRUST by L. Dudley Stamp (Harrap, 1951), pp. 120, 32 coloured plates, 18s.

Described as a new approach to physical geography and geology, the substance of the book is provided by a series of models, photographed in colour and monochrome, with the text consisting of little more than captions expanded. There is a differentiation in treatment between models presented as accurate in detail and others in which some degree of generalization has been effected; the underlying geology being shown diagrammatically in the latter.

ABOUT OUR CONTRIBUTORS

C. L. BOLTZ is qualified academically in physics, mathematics, and psychology. He is a Fellow of the Physical Society and a member of the Society's Colour Group (on committee 1948–50). He is the author of several books – popularly written, simple texts – on radio, electricity, and physics, a subject which he taught for some years before devoting himself to full-time writing.

J. A. V. BUTLER, who has been a lecturer at Edinburgh University, a Rockefeller and a Courtauld Fellow, is now in charge of research in biophysical chemistry in the Chester Beatty Research Institute, the Royal Cancer Hospital, London. Among his publications are a textbook on *Chemical Thermodynamics*, and a popular account of the nature of life *Man is a Microcosm*. He is also joint editor of *Progress in Biophysics and Biophysical Chemistry*.

W. F. FLOYD is Senior Lecturer in Physiology at The Middlesex Hospital Medical School, London. He has been engaged in electro-physiological research since 1934, and is co-author of a recent book *Electrical Phenomena at Interfaces*. He is a member of the Joint Committee on Lighting and Vision of the Medical Research Council and Building Research Board, and is joint honorary secretary of the Ergonomics Research Society.

L. L. GREEN graduated at King's College, Cambridge, and did research work and took his Ph.D. at the Cavendish Laboratory. He is at present a lecturer at Liverpool University, where he works in the Nuclear Physics Research Laboratory.

R. W. JAMES was until lately a member of the research staff of the Section of Meteorological Physics, the Commonwealth Scientific and Industrial Research Organization, Australia, where he worked on the dynamics of cyclones.

F. KINGDON-WARD is a Cambridge graduate who has spent the greater part of his life in plant-collection and exploration in south-east Asia, notably in the high mountain country of western China, north-eastern Burma, Assam and south-eastern Tibet. He was awarded the Royal Medal of the Royal Geographical Society in 1930. He is the author of a number of books including *Modern Exploration*.

[*Continued overleaf*

P. H. S. SILVER is Senior Demonstrator in Anatomy at The Middlesex Hospital Medical School, London, and has recently done research work in functional anatomy on movement and posture.

H. LEES was born in 1917 in Leeds. Honours graduate in biochemistry (Liverpool), Ph.D. (London); worked in industry for three years before going to Rothamsted in 1942, where he became interested in problems connected with plants and soils. Later, he worked for two years in Trinidad before joining the staff of Aberdeen University as Lecturer in Biological Chemistry.

2

(2)	(1)	(2)

SCIENCE NEWS

Outstanding articles from recent issues include:

SCIENCE NEWS 18

The Concept of National Character	*Geoffrey Gorer*
A Hundred Years of Thermodynamics	*J. A. V. Butler*
Determining Atomic Patterns	*Prof. Kathleen Lonsdale*, F.R.S.
Plant Hormones and the Control of Flowering	*Joyce A. Bentley*
Theoretical Implications of Telepathy	*Margaret Knight*
The Scientific Status of Psychology	*Michael Scriven*

SCIENCE NEWS 19

Astronomy Today	*David S. Evans*
Guidance Devices for the Blind	*Franklin S. Cooper*
Fifty Years of Physics	*Prof. Max Born*, F.R.S.
Elusive Chemicals	*George Porter*
Lightning and the Spark Discharge	*R. D. Connor*
Hormones and Carbohydrate Metabolism	*E. Reid*

SCIENCE NEWS 20

The Shape of Inorganic Molecules	*R. S. Nyholm*
Colours of the Stars	*William H. Marshall*
Advances in Anaesthesia	*V. Keating*
Photogeology	
The Origin of Language	*R. A. S. Paget*
The Science of Lubricating Greases	*T. D. Smith*

SCIENCE NEWS 21

Present Concept of the Physical World	*H. J. Bhabha, F.R.S.*
Radio Astronomy	*A. J. Higgs*
Identical Twins in Dairy Cattle Improvement Investigations	*John Hancock*
Forest Entomology	*J. J. de Gryse*
Theory of Crystal Growth	*W. K. Burton*
The Cosmological Problem	*G. C. McVittie*