21

Radio Astronomy
Identical Twin Dairy Cattle
Forest Entomology
Concept of the Physical World
Theory of Crystal Growth
The Cosmological Problems
Research Report

2/-

First published August 1951

*

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

Made and printed in Great Britain
for Penguin Books Ltd
by C. Nicholls & Company Ltd

CONTENTS

EDITORIAL

IN a journal such as *Science News*, there is a pressure towards parochialism in the choice of authors which is hard to resist. Articles are more often offered, and discussion and adjustment are easier, in the case of those authors who are most accessible. There have been exceptions, admittedly, in earlier issues: Maurice Oudot has written on eye-grafting, Sripati Chandrasekhar on the population problems of India and Pakistan, and there have been distinguished contributions from the United States. It may be well, none the less, that there should be a deliberate widening from time to time, and in the present issue there are four articles from authors in Commonwealth countries. There has been no attempt at completeness: it is not intended to present Commonwealth science as something distinct in itself, nor to give any systematic account of work in progress. The aim, rather, has been to include a number of subjects on which Commonwealth scientists happen to be natural contributors, while not excluding other subjects thought to call for early attention. Dr W. K. Burton, for example, writes on a theory of crystal-growth with which he has been associated, and which has lately received further confirmation; and Professor McVittie illustrates the observational approach to cosmology, as distinct from the procedure of 'making guesses as to the nature of the universe', and so reaching conclusions which, having the appearance of certainty, 'would be regarded by many astronomers as more than debatable'.

Of our Commonwealth contributors, Dr H. J. Bhabha is a theoretical physicist whose views command universal attention, as much from the clarity of his thought as from his personal achievement, though this has been considerable. His article on 'The Present Concept of the Physical World' is an adaptation of his recent Presidential Address to the Indian Science Congress. Its first half consists of an analysis of the nature and methods of theoretical physics, with illustrations from the work of Newton and Einstein: in the second half, Dr Bhabha discusses some recent and contemporary problems, ending with some new ideas on the relationship between the mass and the charge of elementary particles. It would be difficult to imagine any greater contrast than there is between this article and the account by John Hancock of an experiment in the use of identical-twin dairy cattle. The latter is of interest, not only in relation to dairy-cattle improvement, which is of particular importance to the economy of New Zealand, but also as the first stage in the largest and most thorough investigation yet undertaken of the respective contributions made by nature and nurture under defined conditions. His account provides its own caution on the unwisdom of too-easy generalization. There is a similar connexion between Canadian resources and the subject of forest entomology, on which J. J. de Gryse is a recognized authority as well as the responsible administrator on the side of research. Finally, from Australia, there is an article by A. J. Higgs on radio-astronomy, on which the Division of Radiophysics of the Commonwealth Scientific and Industrial Research Organization has made important contributions.

It is the general practice of journals that the articles of contributors should be preceded (or not) by an editorial, as a uniform custom; and this uniformity has been followed in *Science News* up to the present time. It appears a possible distinction that, whereas there are occasions when an editorial serves some useful purpose, there may be others when there is no further reason for its existence than that it is normally provided. Should the second happen in future, there will be no editorial.

THE PRESENT CONCEPT OF THE PHYSICAL WORLD

H. J. BHABHA, F.R.S.

THE multitude and variety of the phenomena of nature, which still fill us with astonishment, must have bewildered and awed primitive man. It is not strange that he should have sought, on the one hand, to gain some control over them by investing them with anthropomorphic personality which could be influenced by entreaty and prayer, and on the other to alter his immediate physical environment so as to provide some little shelter or margin of safety against the more hostile acts of nature. This urge eventually led to the early civilizations and the later developments following from them. These civilizations depended on a considerable body of practical knowledge acquired empirically, and some highly developed arts and crafts. A few crucial inventions such as that of the horse harness in China, or of the zero in mathematics in India, had a profound influence on their historical development. But with a few notable exceptions scientific activity in the modern sense did not begin till the Italian renaissance.

Towards the end of the fifteenth century Leonardo da Vinci wrote in one of his manuscripts which is now in the library of the Institut de France (G 96 v):

> There is no certainty where one can neither apply any of the mathematical sciences nor any of those which are based on the mathematical sciences.

This was not the mere expression of a specialist extolling his own subject: it was written by perhaps the most versatile genius the world has known, one who had a greater mastery of all the various arts and sciences of his time than anyone since. It expresses the new spirit of the times, a spirit which was to lead eventually to that vast development which is modern science and technology. 'The mathematical sciences' for Leonardo consisted of what had been handed down of Greek mathematics, while by the sciences based on the mathematical sciences he understood the applications of geometry to optics and mechanics. What Leonardo wished to emphasize, I feel, was that as long as an observation of a natural phenomenon remained couched in qualitative terms it would not be definite enough to build on, and only by introducing accurate measurement and quantitative relations into it could one be certain that it was right or wrong within the limits of accuracy of the measurements. Some four centuries later, Lord Kelvin was to say:

> When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.

Once this general approach received fairly wide acceptance, the development of science in the modern sense was inevitable.

It was found quite soon that certain properties, which could be stated in terms of exact measurement, were common to many objects. In certain cases, therefore, it became possible to state a general property without specifying the particular object to which it belonged. Such general properties could then be regarded as laws or regularities of nature which all objects of a certain type satisfy. One such regularity or law of nature was the one discovered by Archimedes, that the loss in weight of a body immersed in water is equal to the weight of the water displaced. Archimedes was indeed one of the shining forerunners of modern science. It is nevertheless interesting to note that his law is a law in statics. Laws involving the motions of objects were to come much later. An example of a dynamical law is the regularity discovered by Galileo, that all heavy bodies fall the same distance under gravity in a given interval of time irrespective of their weight. Other regularities of the same type, but which involve more complicated relations between the objects, are the three laws of Kepler on the motion of the planets.

Laws or regularities of nature of the type just mentioned are merely empirical statements of properties observed to be common

to a large number of objects. They are all unconnected with each other. In order to connect up such regularities, it may be necessary to formulate certain more abstract principles or postulates from which the various observed regularities can be deduced.

Newton's fundamental laws of motion exemplify this new approach. Consider his first law, which reads:

> Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by a force impressed upon it.

Assuming that we understand intuitively what is meant by rest or uniform motion in a right line, we may well ask ourselves what is meant by an impressed force. If we turn to the definitions which Newton has placed a few pages earlier at the beginning of his *Principia*, we find the answer in Definition IV:

> An impressed force is an action exerted upon a body, in order to change its state, either of rest, or of uniform motion in a right line.

Expressed in this way Newton's first law would appear to be a tautology: it states that a condition A exists unless interfered with by the existence of B: while B is defined to be present when the condition A is interfered with. It is not the purpose of this discussion to minimize in any way Newton's achievement, which is one of the greatest monuments in the history of science, but to understand the real nature of his laws. Let us assume that force can be defined in some other way than in the above definition so as not to make the first law a tautology. One might then imagine the first law to be a statement arrived at from direct observation through some process of induction. For example, in commenting on his first law, Newton writes:

> Projectiles continue in their motions, so far as they are not retarded by the resistance of the air, or impelled downwards by the force of gravity. A top, whose parts by their cohesion are continually drawn aside from rectilinear motions, does not cease its rotation, otherwise than as it is retarded by the air. The greater bodies of the planets and comets, meeting with less resistance in freer spaces, preserve their motions both progressive and circular for a much longer time.

The inference is that one may conclude by induction that if we could take a body into space to a very great distance from all other material bodies then it would either remain at rest or move in uniform motion in a straight line. We know to-day that such an induction cannot be made, and may indeed not even be true for the actual world. While it is possible mathematically to assume a world in which Newton's laws are strictly true, it is equally possible mathematically to think of worlds in which they are not. This analysis shows us that strictly speaking *Newton's laws of motion and gravitation are abstract mathematical statements which he quite rightly calls axioms.* And if they came to be regarded as objectively true it is because the behaviour of objects which could be deduced from them by mathematical reasoning agreed with our direct observations. For example, one could deduce from Newton's laws the regularity observed by Galileo concerning the fall of bodies, the three regularities observed by Kepler on the motion of the planets, and a host of other phenomena. Quite appropriately, his epoch-making work was called 'The Mathematical Principles of Natural Philosophy'.

The great importance of the contribution of Newton to the development of physics is that it introduced a new approach into science. It led to the acceptance of the position that the ideas which are to be regarded as fundamental for the understanding of nature are certain abstract concepts or postulates which cannot be proved directly, and not the directly observable regularities of nature which can be deduced from them. This position was accepted because it allows one to order different empirically found regularities of nature into a unified logical scheme which would not otherwise be possible.

A consequence of this approach is that any newly discovered fact of nature which does not fit into the existing scheme of physics may necessitate a complete change of the fundamental postulates. Since, however, the old postulates were such that a very large body of observed facts about nature could be deduced from them, it follows that they must still have a restricted validity under certain circumstances, and be deducible as approximations from the new postulates. Although, therefore, every new

discovery which does not fit into the old scheme necessitates a complete change of the fundamental postulates, the change is always from a more limited set of concepts to a set of more general concepts. As one goes deeper into the understanding of nature by co-ordinating all the known facts into one scheme by the use of wider concepts as the basic postulates, the old fundamental postulates become, in a sense, a part of the superstructure, taking a place in between the new fundamental concepts and the directly observed regularities of nature.

An example of this process of generalization of the basic concepts is the well-known development from the pre-relativity concept of an absolute space and an absolute time to the more general concept of the unified space-time of the theory of relativity. The example shows how the basic concepts of a theory may be radically changed, while still retaining most of the notions of the earlier theory, but recognizing them to be of limited validity, true not universally but only in certain circumstances. Thus, the absolute distinction between a time interval and a space interval in pre-relativity physics is replaced in relativity theory by the absolute distinction between time-like and space-like intervals, while the notion of the absoluteness of time of the earlier theory is seen to be approximately correct in the new theory for a group of observers moving relative to each other with velocities small compared with that of light. The example shows also that a widening of the basic concepts automatically reduces the amount of arbitrariness in the theory. In pre-relativity theory the force between two bodies could be taken to be entirely arbitrary, so far as theoretical requirements were concerned. In relativity theory, on the contrary, the force has necessarily to be conveyed through the medium of a field. The basic differential equations which any field has to satisfy in relativistic theory are drastically restricted in their variety by the requirement that the laws of nature should be so formulated as to have the same form for all observers moving relative to each other with uniform velocity. With this restriction, there is a very limited freedom in choosing the form of the force which can be exerted between two particles.

When a science reaches an advanced stage, as physics undoubtedly has, the facts which can be discovered by direct observation become more and more meagre. We may expect, for example, to be able to discover by experiment the masses of the various types of elementary particles, the different processes which they undergo, their general behaviour in passing through substances of different types and so on. It is inconceivable, however, that an equation like the equation for the electron which Dirac put forward in 1928 could be deduced by direct observation in the sense that Maxwell deduced the equations of the electromagnetic field from observations made by Faraday. In such an advanced stage, however, there are compensating advantages. We have a number of theories to fall back upon with the knowledge that each correctly describes a large body of experimental facts in certain circumstances. We can, therefore, attempt to proceed by evolving new theories which reduce to the previously known ones in the circumstances in which the latter are known to be correct. There is, indeed, no other course open to us.

The aim of theoretical physics must be to find a complete set of mutually consistent mathematical postulates or axioms from which the properties of nature, defined as the result of every conceivable experiment, can be deduced in the form of a number of theorems. In order, however, to compare the mathematical statements of the theorems with the results of observation, the basic mathematical postulates must be supplemented by a set of prescriptions about the interpretation of the mathematical formalism. It is clearly not sufficient that the postulates should be consistent. Their correctness from the point of view of physics can be demonstrated only by an agreement between the deductions and the results of experiment.

It is most important to distinguish this approach from the one which assumes that one can arrive at the laws of nature by pure thought and epistemological reasoning. The latter approach has neither met with much success, nor proved particularly fruitful in promoting an understanding of the physical world. In the approach described on the contrary, we recognize that it may be possible to build many logically consistent theories which have nevertheless nothing to do with the actual structure of the physical world. Theories in pure mathematics provide many such examples. If any set of axioms or postulates can claim to correspond with reality it is because the deductions from them stand the test of agreeing with the results of experiment.

We can turn now to review the development of our picture of the physical world resulting from recent discoveries. It had already been established by the end of the last century that the multitude of substances in nature are all made up by the chemical combinations of a certain number of basic substances called the chemical elements. The smallest unit of a given chemical element was called an atom. The combinations of these atoms, either of the same element or of different elements, gives rise to chemical compounds, which compose the body of all the substances that we meet in nature.

Investigations on the conduction of electricity through gases led Thomson towards the end of the last century to the discovery of the fact that this conduction could be attributed to a particle of negative charge having always the same ratio of charge to mass irrespective of the substance under investigation. Moreover, this mass was some thousand times smaller than that of a positive ion. Subsequent researches have established the fact that there is a smallest unit of negative electricity, smaller sub-divisions of it not being found in nature, and that all negative electricity appears in integral multiples of this smallest unit, which is now denoted by e. Thus it came to be established that there was a type of particle, called an electron, which always possessed the same negative charge e and the same mass m, which was somehow contained in atoms, and whose behaviour was responsible for the phenomenon of electricity.

Since an atom is an electrically neutral body, it follows that if electrons are contained in it, then it must also contain an equal amount of positive charge. It was not clear at the time how the electrons and the positive charges of electricity were distributed in an atom. For example, were the electrons embedded in a uniform medium of positive electricity, rather like plums in a cake? Or were they like planets revolving round a sun of positive charge? The answer to this important question was furnished by Rutherford in 1911. He showed by a study of the scattering of alpha-particles that the true picture of the atom was to consider it as like a solar system in which the electrons move like planets round a heavy centre called the nucleus in which all the positive charge and most of the mass of the atom is concentrated. Since the negative charge inside the atom depended on the number of electrons in it and was an integral multiple of e, the positive charge on the nucleus had likewise to be an integral multiple of e. It was soon established that the number of units of positive charge on the nucleus determined the chemical properties of the atom, and that there were 92 such chemical elements ranging from the lightest, hydrogen, to the heaviest, uranium, with 1 to 92 positive units of charge on the nucleus respectively. The mass of the nucleus of the lightest element, hydrogen, containing just one unit of positive electricity, was found to be always precisely the same, and some 1840 times the mass of the electron. Since this nucleus of hydrogen never broke up into smaller fragments, it became convenient to regard it as a new type of fundamental entity, a new elementary particle, called a proton.

Further researches showed that the mass of any atom was always almost precisely an integral multiple of the mass of the proton, while its charge was a smaller integral multiple of the charge of the proton. These facts led one at the time to accept a picture of the nucleus which made it appear to be made up of protons and electrons only. The number of protons was sufficient to make up the mass of the nucleus, while a certain number of electrons were added inside the nucleus to neutralize the charge of some of the protons and make the total positive charge equal to the actual charge of the particular nucleus. Thus round about 1930, our picture of the physical world appeared to be remarkably simple. The whole material world was thought of as made up of just two types of elementary particles, protons and electrons. By suitable arrangements of these, one built up the atoms of the chemical elements; and from suitable arrangements of the latter, every other material thing that was found in nature. Light, or, in more general terms, electromagnetic radiation or photons, and gravitation, were the only two other physical entities found in nature.

A scientist at that time could have thought, as many did think, that when one knew the mathematical laws governing the behaviour of these four elementary types of physical entities, the protons, electrons, photons and gravitation, one would know everything of a fundamental nature that there was to know of the physical world, and physics in principle would be a subject which had reached its destination. The developments of the last twenty years show us how far from the truth was this belief. It shows in a striking manner that however great the successes of a theory, unless this success is complete and total, it is always possible that something very important may have slipped through the net. The apparently small but persistent difficulties or inconsistencies in a theory, or small discrepancies between theory and observation, may be essentially unbridgeable within the framework of the basic concepts of that theory and yield the clue to new ideas.

What were these difficulties? There was an inherent contradiction which arose from the inclusion of electrons within the nucleus. To understand this contradiction, we have to consider the property of 'spin'. In order to make a body spin about itself like a top, we have to impart to it energy and something called angular momentum. It is found that like electricity, which occurs in nature only in integral multiples of a basic unit e, so angular momentum also occurs only in integral multiples of a basic unit. This unit involves Planck's constant (h), although Planck's constant was introduced in the first instance as a constant affecting radiation rather than angular momentum. The basic unit of angular momentum is Planck's constant divided by 4π, that is $h/4\pi$. Experimentally, we can determine the angular momentum of the proton and electron by spectroscopic analysis. In this way, we know that each possesses an angular momentum or spin of one unit; this arises, so to speak, from its spinning about itself like a top. On the other hand, angular momentum which arises from one body moving bodily round another, is always an even multiple of the basic unit, that is either zero or two or four times, etc., $h/4\pi$. If the nucleus were made up of protons and electrons we should expect both to contribute to its angular momentum. But on the contrary, we find from spectrographic analysis that the spin of a nucleus is an odd or an even multiple of the basic unit, depending on whether the number of heavy particles in it (protons in our picture) is odd or even, irrespective of the number of electrons supposed to be in it. It was as if the electrons in the nucleus made no contribution to the spin, although every electron outside the nucleus contributes an odd number of units, due to its intrinsic spin and its bodily motion.

Secondly, both protons and electrons satisfy a law which the theoretical physicist calls Fermi-Dirac statistics. It can be shown then that a nucleus containing an odd number of protons plus electrons must also satisfy the same statistics so that in a molecule composed of two such atoms only certain spectral lines must appear and not others. Experiment again shows that the statistics of such nuclei appear to depend only on the number of heavy particles in the nucleus and not on the total number of protons plus electrons in our picture. In fact all the nuclei seem to behave as if the electrons which were supposed to be in them only manifested their electric charge but neither their spin nor their statistics. A bold attempt to face this difficulty would soon have led one to the view that nuclei were not composed of protons and electrons but rather of protons and some hitherto unknown particle having to a very high degree the same mass as the proton, the same spin and satisfying the same statistics. A particle of this description was discovered by Chadwick in 1931 and was called a neutron. It had to be accepted as a new elementary particle and not a composite structure made up of a proton and an electron for the same reason that prevented us from thinking consistently of the nucleus as being made up of protons and electrons. It immediately led to the acceptance of the picture that all nuclei are composed of only two types of particles, protons and neutrons. The number of types of elementary particles was thus increased by one.

The acceptance of the neutron as an elementary particle, however, introduced a new feature into our concept of the elementary

particles. For it had been known for a long time that certain nuclei, as for example those of the radioactive elements, emit electrons every now and then. One can fit this fact into the picture only by assuming that when such an electron is emitted from the nucleus it is in fact newly created in the process and that simultaneously a neutron in the nucleus changes into a proton. Thus we have to admit the possibility that while the elementary particles are not composite, and that as long as they exist they are immutable with absolutely constant properties, nevertheless there are occasions when one or more such particles can disappear altogether with the simultaneous creation of another set. For example, a neutron may disappear and give place to a proton and an electron. Since the neutron, proton and electron all have a spin $h/4\pi$, and the bodily motion of these particles can only contribute an even multiple of $h/4\pi$, the conservation of angular momentum and statistical properties compels us to postulate that there must be yet another elementary particle, called a neutrino by Pauli, which possesses no charge, a mass negligible compared with that of the electron, and a spin of one unit ($h/4\pi$).

In 1931, Anderson, in California, reported a photograph which seemed to be that of a particle of the same mass as an electron and having one unit of positive instead of negative electric charge. The experimental advance of making the cosmic rays themselves take their own photographs instead of taking photographs at random in a Wilson cloud chamber then enabled Blackett and Occhialini, in Cambridge, soon afterwards to discover a new phenomenon called cosmic ray showers. Although cosmic rays are a relatively rare event, Blackett and Occhialini showed that very frequently many such rays occurred in a shower, and subsequent work has demonstrated that such showers of particles are produced by cosmic rays when they pass through matter, as for example sheets of lead placed in the Wilson chamber. Blackett and Occhialini showed that their showers contained not only the usual electrons but a comparable number of electrons with the opposite charge. With this, the existence of the positron, as this new particle was called, was established.

The existence of the positron could be understood immediately in terms of an equation for the electron which Dirac had put forward in 1928 and which combined in it for the first time the ideas underlying the theories of relativity and quantum mechanics. Dirac had already shown that certain apparent difficulties in his theory could be understood as expressing on the one hand the existence of a particle of equal but opposite charge to that of the electron and on the other the possibility of a pair of such positive and negative particles being created by the materialization of energy, or of their annihilation with the transformation of their mass-energy into radiation. Subsequent experiments have fully confirmed the correctness of these basic processes predicted by the theory. Nevertheless, a consequence of this theory was that no electron or photon of even the highest energy could penetrate large amounts of matter, while a growing body of evidence from cosmic ray experiments indicated that particles which looked like electrons did in fact penetrate great thicknesses of matter. Thus, there seemed to be evidence that quantum theory failed for very high energy electrons, while at the same time there was no theory to explain the phenomenon of the cosmic ray showers. It was shown by Heitler and the present writer that the existence of cosmic ray showers, and the behaviour of the soft component of cosmic rays in the atmosphere and in dense substances, could be explained on the basis of quantum theory. One initial particle of high energy gives rise to a progressively greater number of particles of lesser energy, like a cascade, until finally the whole energy is dissipated in the substance traversed. Only after the cascade theory had been formulated could it be concluded that the electron-like tracks of particles, which did not behave completely like electrons nor like protons, must be due to a new type of particle of intermediate mass. Thus the existence of a new particle called the meson, with a mass some 210 times that of the electron, came to be established in 1938.

A particle with a mass of this order of magnitude had already been envisaged by Yukawa in 1935 in an attempt to explain the short-range nature of nuclear forces, that is the forces between two particles in the nucleus of an atom, as for example a proton and a neutron. The observed mesons therefore came to be regarded as the agency responsible for nuclear forces; and radioactivity characterized by the emission of an electron was considered as due to the decay of virtual mesons emitted by nuclei. Research carried out since the end of the war has demonstrated that the picture was, again, not so simple as it was supposed to be.

Firstly, although the decay of mesons into electrons has been confirmed by experiment, more accurate experiments have shown that the electrons are emitted with a continuous distribution of energies, and not with a single, sharply defined energy, as originally supposed. There must, in any case, be at least one neutral particle emitted, as well as the electron, to secure a balance of energy and momentum. From the continuous distribution of energies of the electrons, it can be concluded that not one, but two neutral particles are emitted simultaneously.

Secondly, the identification of the observed cosmic ray mesons with the particles responsible for nuclear forces inevitably requires them to have a strong interaction with nuclei, whereas the mesons observed in cosmic rays were seen to penetrate large amounts of matter with but a very weak interaction with nuclei. A serious attempt to face this difficulty might easily have led one to the conclusion that the particles responsible for nuclear forces were not in fact the observed mesons. This conclusion was, however, not accepted until the studies by Powell and his group, at Bristol, of the tracks of cosmic ray particles in special photographic plates had shown that there are in fact two types of mesons with two different masses, and that the one type decays into the other with a period of about one hundred millionth part of a second. The two types of mesons are now known as π- and μ-mesons respectively. The meson generally seen in cloud chamber photographs are the μ-mesons, whereas the π-mesons are the ones which are now identified as being responsible for nuclear forces. In accordance with this picture one would have to attribute an even integral spin (in units of $h/4\pi$) to π-mesons.* The mass of π-mesons, or pions as they are called in short, as determined from measurements

* There is no connection between the π symbol, used for the designation of π-mesons, and the π of mathematics, which appears in the expression for units of spin.—ED.

of their deflection in a magnetic field, appears to be in the neighbourhood of 276 times that of the electron.

More recent experiments with the large cyclotron at Berkeley, in California, and elsewhere have led to the discovery of yet another new elementary particle, a neutral pion, that is a pion with practically the same mass as the charged pion but with no electric charge. Although such a neutral pion interacts strongly with nuclei, nevertheless it cannot be observed directly in a photographic plate or in a cloud chamber due to the fact that it does not possess an electric charge and therefore does not ionize. Despite this, however, it has been possible to ascertain its mass with very great accuracy (see *Science News* 18, pp. 88-90). It is found that the mass of a neutral pion is only a few electron-masses less than that of a charged pion.

Rossi has shown that in cosmic rays some two-thirds of the total energy is converted into charged penetrating particles while one-third disappears into neutral charged particles. Since charged pions can have a positive or a negative charge, it follows from this that the interactions of a positive, a negative or a neutral pion with a nucleon are roughly of the same magnitude.

I now come to particles whose existence is highly probable, though not absolutely certain. In 1947 Rochester and Butler in Blackett's laboratory in Manchester reported certain unusual events which they had observed in a Wilson chamber. They occasionally saw the tracks of charged particles which seemed to show an abrupt change in their direction in the gas of the chamber. They showed that it was difficult to interpret these tracks as due to scattering in the gas for two reasons. Firstly no recoil nucleus in the gas was visible. Secondly it would be difficult to understand why a particle should have such a great likelihood of collision while passing through the rarified medium of a gas and yet pass through dense matter like a lead plate without suffering any collision at all. They put forward the explanation that these forks were due to the spontaneous decay of a charged particle into another charged particle and a neutral one. Experiments at present being carried out at Pic-du-Midi and Jungfraujoch are rapidly producing further evidence that this interpretation is

correct. It would be consistent with the present evidence to interpret the charged particle resulting from the decay as a pion. Whether all the original particles are of the same type is not a question that can be answered at present. Most of the particles appear to have a mass in the neighbourhood of 800 times the electron mass. But there is some indication that there may also be particles of this type with masses more than a thousand times that of the electron. Rochester and Butler, and more recently Butler and his collaborators, have also produced evidence to show that neutral particles of corresponding mass exist which seem to decay into two charged particles in the same way. Lastly one should mention the case of a particle observed by Powell and his group and called by him a τ meson, which came to the end of its range and emitted three mesons of which one is certainly a pion.

We see now that at least nine different types of elementary physical entities exist in nature, while the existence of two more is almost certain. While experiments may give us information about the masses of these particles, their mutual interactions and the processes in which they take part, it seems inconceivable that an experiment would enable us to deduce directly the mathematical equation describing the behaviour of any such particle. We can only hope to set up the mathematical equations governing the behaviour of these particles by taking as our guides certain well known principles, as for example the principle of relativity and the ideas underlying quantum mechanics. Even such a clearly defined property as the spin of an elementary particle is not something we can hope to measure directly in the case of particles like the meson, but must infer it from considerations of the processes in which they take part, by comparing the behaviour of particles of different spins as predicted by theory with the experimental observations.

The circumstance that there are a dozen different types of elementary particles in nature would lead us to expect that there may be many more, and indeed with our present knowledge we cannot exclude the possibility that there may be an infinite number of them. This does not mean, however, that we shall never be able to obtain a complete description of them all. There are, for example, an infinity of lines in the spectrum of hydrogen and yet we possess today not only a formula which in one neat expression contains the energies of all these lines but also a mathematical theory which allows us to calculate all the stable states of the hydrogen atom, other properties such as the scattering of electrons by atoms, their creation by photons, and even more complicated properties like the nature of the chemical bond between two hydrogen atoms. It is, therefore, quite possible that with increasing knowledge we may be able to find the formula which gives us the masses of all the elementary particles and the general principles which will allow us to deduce the equation satisfied by a particle of any particular mass.

Lorentz at the beginning of this century regarded the charge of an electron as a property of the electron and tried to explain its mass as due to the energy of the electro magnetic field associated with that charge. The idea that the mass of the elementary particle is wholly of field origin has had to be abandoned because we know a number of elementary particles all having the same electrical charge but different masses. The empirical situation which faces us leads on the contrary to a different conclusion. We know that an electromagnetic field, in the form of freely moving quanta of radiation, exists in its own right, quite apart from the sources from which the quanta were emitted. On the other hand, whenever the quanta interact with other types of particles, the measure of this interaction, which we call the charge, is always the same whether the particle is an electron, a meson or a proton. From a phenomenological point of view, therefore, we would be more justified today in considering the electric charge e of an elementary particle as a property of the electromagnetic field rather than of the particle, while considering the mass of the particle as an intrinsic property of the particle, unconnected with its interaction with the electromagnetic field. Our approach to this problem today should therefore be just the opposite of that of Lorentz. If the charge e is to be considered as a property of the electromagnetic field, as here suggested, then since the only unit associated with the field in which it could be measured is the square root of Planck's constant multiplied by the velocity of light, we should consider

this ratio, or its square e^2/hc to be an intrinsic property of the electromagnetic field. The dimensionless constant e^2/hc, which has the value 1/137 in nature, would then appear to be a number associated with the electromagnetic field, and not a universal constant of nature of the same status as Planck's constant h or the velocity of light c which enter into the description of other elementary particles.

It is clear that we are now penetrating into a new level of nature which was practically unknown some twenty years ago. I have pointed out earlier that although there may be an infinite number of types of elementary particles nevertheless this fact in itself does not necessarily force the conclusion that we will never be able to describe nature fully or to explain the physical world exhaustively. On the other hand we cannot be certain with our present knowledge that a complete mathematical theory of the physical world can be based upon a finite number of postulates, and if this were not so we would be faced with a situation in which we could never hope to give an exhaustive description of everything there is in nature, but only to extend with the flow of time the region which we had explored and understood.

REFERENCES

Bhabha, H. J. *Reviews of Modern Physics*, **21**, 451 (1949).

Cajori, F. *Newton's Principia: a revision of Motte's translation.* Berkeley, University of California Press.

MacCurdy, E. *The Notebooks of Leonardo da Vinci.* London, Jonathan Cape, 1945.

[*Note by Editor.* Experimental work on meson particles, mentioned in the second half of this article, was reviewed by Terence Price in *Science News* 16; and again, under 'Research Report', in *Science News* 18. Of instruments and experimental techniques, here mentioned in passing, the Wilson cloud chamber was described in *Science News* **2**, 130; the use of the photographic plate as a research instrument in **5**, 56; and the cyclotron in **2**, 149.]

This article is adapted from the Presidential Address to the Indian Science Congress delivered on 2 January 1951 at Bangalore.

THEORY OF CRYSTAL GROWTH

W. K BURTON

In spite of the great interest aroused by the crystalline state, it is only recently that an understanding has been obtained of how crystals grow. The interpretation of a crystal as a regular arrangement or *lattice* of ions, atoms or molecules was one of the early triumphs of the atomic theory, and this interpretation has been confirmed in a variety of ways. Probably the most important experimental verification of the existence of the regular lattice is the phenomenon of x-ray reflexion by crystals. When reflexion takes place, it is as if there were certain reflecting sheets in the crystal, which are interpreted as planes of atoms. However, reflexion does not take place at *any* angle between the x-ray beam and the plane of atoms concerned, but only for those angles which are related in a certain way to the wave length of the x rays and the spacing between the atoms in the plane. This fact enables one to identify the various reflected beams as having arisen from particular planes of atoms in the crystal, which makes it possible to construct a picture of how the atoms are arranged. (*See* article by Kathleen Lonsdale in *Science News* 18.) The dependence of the possible angles of reflexion on atomic spacing, for a given x-ray wavelength, enables one to measure the atomic spacing and thus to determine the scale of the crystal. Once this has been done, a measurement of the external dimensions of the crystal gives the number of atoms contained in it, and weighing the crystal then leads to a knowledge of the mass per atom of the substance composing the crystal. The value obtained agrees with the results of quite different determinations of the mass per atom. So the hypothesis that crystals consist of regular lattices of atoms appears to be very well founded. There is, however, one discrepancy, the significance of which

will be seen later. It is that the intensity of the reflected x rays is much higher than the theory predicts. This must mean that the crystal, although exceedingly regular, is not perfectly so: there must be slight distortions of the lattice which relax the condition for reflexion through small variations either in atomic spacing or in the flatness of the atomic planes. These departures from the ideal lattice need only be very slight to account for the superior reflecting power of actual compared with ideal crystals.

THE GROWTH OF IDEAL CRYSTALS

In order to formulate a theory of growth, it is necessary to choose a *model* for a crystal. By a model is meant a simplified representation of a real crystal in which certain features thought to be irrelevant to the problem under discussion are omitted for ease of calculation. If the properties of the model turn out to be different from those of the real crystal, this will show that the model is unsuitable, and it must be modified. The modified model may represent the known facts quite well, and in this case one looks for further experimental facts, suggested by the model itself, to see if the new model really does contribute in a significant way to our understanding of real crystals. A model can never be *proved* to be 'correct'; it can be shown only to be either consistent or inconsistent with observation. The aim is, of course, to obtain as realistic models as possible. This interplay between theory and experiment is well illustrated by the development of the theory of crystal growth.

Let us now commit ourselves to a definite model for a molecular crystal, such as iodine. A plausible model for this case is the following, due to Kossel: the molecules are cubes, and are stacked face to face. Each cube is attracted equally by all its six neighbours and only nearest neighbours attract one another. The kind of packing just described is called simple cubic packing. As far as is known, no crystal with this structure exists in nature, but the results for other structures are not essentially different. The molecules are represented by cubes instead of spheres or some other shape for ease of picturization: the molecules are assumed to lie only within the cubes. The energy required to separate two cubes is called the binding energy, and in general is different for different materials.

The question is now asked: how does such a crystal grow from a gas or vapour? Let us consider first the case of the further growth of a crystal already assumed to be present, leaving until later the question of how the crystal arose in the first place. It is convenient to start by considering the arrangement of molecules at the 'absolute zero' of temperature, at which there are no thermal vibrations. This temperature, which is the same for all substances, is about 273 degrees centigrade below the freezing point of water. Then unless the number of molecules in the crystal happens to be the exact number to complete all six faces, the appearance of at least one face will be as in Fig. 1. That is, one face will be incomplete, leaving a *step* in the face, and the step itself will be incomplete, giving a kink in the step. In view of their almost certain occurrence on a crystal surface, we can expect that considerable importance attaches to kinks. A new molecule will be 'properly built into' the crystal when it is added to a kink. This addition will move the kink forward by just one position. Conversely for the removal of a molecule.

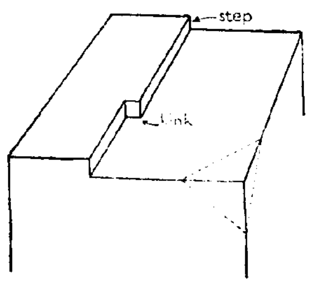


Fig. 1. Crystal surface at –273°C.

Suppose now that the temperature is raised. This will mean that the component molecules will start to vibrate relative to one another, the vibration becoming progressively stronger as the temperature becomes higher. Because of the vibration, molecules will

be shaken from the crystal and will fly about in the space surrounding it. These free molecules constitute the vapour. At first those molecules at kink positions will be more likely to leave the crystal than others, since these are bound by only three neighbours, whilst those in the straight portions of the step have four, and those on the flat portions of surface have five nearest neighbours. But some molecules from every type of position will have some chance of escaping into the vapour. The crystal surface will then look more like Fig. 2. The step will now have many kinks in it, and there will be molecules resting on the surface and against the side of the step: these are called adsorbed molecules. Finally, there will be holes on the surface and in the step.

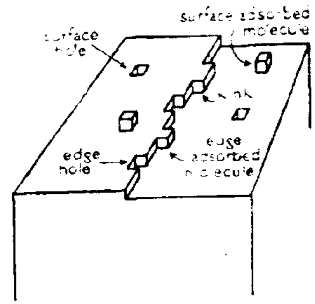


Fig. 2. Crystal surface at moderate temperature

While some molecules are leaving the crystal, others will arrive at the crystal from the vapour. Most of these will become adsorbed surface molecules, a smaller number will become adsorbed on the sides of steps, and a smaller number still will fall directly into kinks. The adsorbed molecules will also be capable of moving over the crystal surface, eventually being absorbed by the kinks. Every process which can occur in one direction can also occur in the reverse direction. Clearly there are an enormous number of possibilities. However, at a certain vapour concentration (average number of vapour molecules per unit volume) as many molecules will leave the crystal in a given time as arrive at it, and no net growth or evaporation will take place. This is called the equilibrium situation; and it can be shown that at this vapour concentration, all the individual processes which take place do so equally in both directions. At a given temperature, the equilibrium concentration of the vapour will be determined by the strength with which the molecules are bound at the kink positions; the more tightly they are bound, the smaller will be this equilibrium concentration. Also the higher the temperature, the greater the vibration, and so the greater the equilibrium concentration of vapour molecules. To work out the equilibrium vapour concentration we first work out the average number of molecules leaving a kink per unit time. At equilibrium this number must be equal to the average number of molecules entering a kink from the vapour, a number which depends on the number of molecules present in the vapour. The number of kinks in a step at equilibrium, and the number of atoms on the surface or on the side of a step can be worked out in a similar way. Typical results, at room temperature, are a kink for every ten molecules along the step, and an adsorbed molecule for every hundred molecules along the step, whilst on the surface there will be one adsorbed molecule per ten thousand positions.

Let us now examine what happens if the vapour concentration is raised above the equilibrium value. In this case the number of molecules leaving a given kink will be the same as before, but the number arriving at it from the vapour will be raised by the same factor as the vapour concentration. The difference represents the net stream of molecules going into the crystal. In other words, the net rate of entry of molecules into a kink is equal to the actual rate at which molecules leave the kink times the supersaturation. The latter quantity is the excess of vapour concentration above the equilibrium value, divided by the equilibrium vapour concentration. If it could be assumed that all the crystal surfaces were entirely covered by kinks, we should have a rate of growth which is the highest possible for a given super-saturation and moreover a rate which is easily calculated. However, there is no reason for believing that kinks can be so plentiful, except in the case of a surface such as is obtained by slicing off a corner of

the crystal (dotted triangle in Fig. 1). For any other kind of surface, we must expect a much smaller number of kinks, and hence a rate of growth which is less than the maximum.

However, surprising as it may seem, experiments on the rate of growth of simple substances show that it is quite usual for a crystal to grow at this maximum rate. Volmer and Schultze published in 1931 the results of some well-controlled experiments on the growth-rates of iodine, phosphorus and naphthalene from their vapours; their figures agree well with those calculated from the above simple theory of maximum growth-rate. Only in the case of iodine is there a significant discrepancy: the rate of growth is less than the expected value when the supersaturation is less than about 1 per cent. The graph in Fig. 3 shows the type of result they obtained, the dotted line being the expected theoretical result.

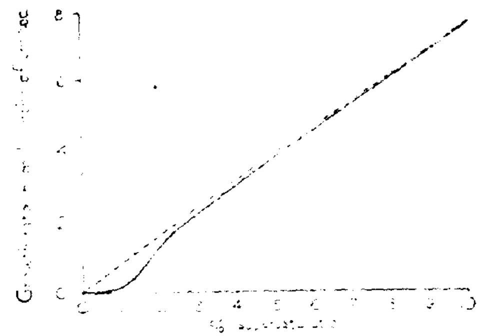


Fig. 3. Dependence of growth-rate on supersaturation for iodine at 0°C.

Although the theory appears to give the right answer, more or less, the situation as so far described is far from satisfactory. We have yet to account for the presence of a large number of kinks on *any* crystal face, not just the faces obtained by chopping off corners. Since in any case the latter type of face would grow more rapidly than any other type of face, such faces would quickly disappear: the crystal would rapidly fill out the chopped-off corner, and growth would then take place by the growth of the slowest-growing faces. It is these faces, then, which we must consider: those parallel to cube faces.

In the case we have been considering so far, there is one step in the face, with kinks spaced at an average distance of about the molecular diameters. Dr Cabrera and I have examined the structure of such a step in great detail, and also the question of how the rate of advance of a step depends on the kink spacing. We found, as might be expected, that the closer the kinks, the faster the step moves, but we found also that the maximum rate is nearly attained as soon as the kinks are closer together than the so-called mean migration distance. This is the average distance a molecule wanders on the surface of a crystal between the time it hits the surface and the time it evaporates again into the vapour. The typical value for this distance is about a hundred molecular diameters, so a typical step can be assumed to grow at its maximum rate. It is, however, incumbent upon us to account for the presence of steps separated by not less than the mean migration distance, *if the theory is to agree with the experimental* results of Volmer and Schultze. If for some reason the steps could become further apart below a supersaturation of 1 per cent, their results with iodine could also be explained. But here everything is against us. For consider what happens to the step (Fig. 2) as it grows: it simply travels to the edge of the crystal and disappears, leaving a *completed* surface containing a few holes and adsorbed molecules. Growth can continue only if a new step can be formed.

It might be thought that perhaps steps can be formed in a similar way to the formation of kinks in a step, but detailed calculation by Dr N. Cabrera and myself shows that a surface like the one we are considering remains perfectly flat apart from the holes and adsorbed molecules unless the temperature is raised almost to the melting point of the crystal, when the surface rapidly becomes rough.

So the steps must be formed gradually, if at all, at ordinary temperatures. A small island of molecules may collect on the surface of the crystal, giving a closed step, which might then grow until the whole surface is covered by a new layer. It turns out that when an island, or nucleus as it is called, is less than a certain critical size depending on the supersaturation, the chances of a new molecule being added are smaller than the chances of a molecule being evaporated from the nucleus. These conflicting circumstances balance at the critical size, and above it the nucleus is more likely to grow further than to evaporate. So we must ask, what are the chances of a nucleus of critical size being formed by the random motions of the molecules on the surface? The answer to this question will give a factor, which when multiplied by the maximum growth rate mentioned earlier will give an upper limit to the growth rate due to the nucleation-mechanism. This factor is, of course, less than one, and in fact, if expressed as a decimal, it turns out to have over a thousand noughts between the decimal point and the first digit different from nought, for a supersaturation of 1 per cent.

This means that the results of the nucleation theory and experiment differ by at least this factor, a discrepancy which must be one of the largest ever recorded. A crystal growing at the theoretical rate would not change perceptibly in size over times as long as the so-called 'age of the universe', say ten thousand million years. A refinement in nucleation theory by Dr Cabrera and myself showed that the number of noughts after the decimal point could be reduced to eight or nine hundred at the best, which is quite useless.

For higher saturations, the number of noughts decreases rapidly, and at a supersaturation of 50 per cent or more, growth could occur at a reasonable rate. There is nothing wrong with the nucleation theory regarded as a set of deductions from premisses. As might be expected, it predicts that there will be an even smaller chance of the initial appearance of a crystal nucleus from a vapour containing no nuclei, at supersaturations of 1 per cent or so. The factor becomes reasonable in this case at supersaturations of a few hundred per cent, and it is in fact observed that such supersaturations are necessary for the production of the initial crystal nuclei.

We must conclude that growth by surface nucleation does not occur at low supersaturations, and hence the Kossel model must be rejected, and we must take a model with an imperfect lattice to represent real crystals.

THE GROWTH OF REAL CRYSTALS

The really significant contribution in this direction came from Dr F. C. Frank, who indicated a reasonable way out of the dilemma. He pointed out that if the crystal contained a certain kind of imperfection, called a dislocation, a step could appear on the surface which would not disappear during growth. Such imperfections had previously been invoked to explain certain mechanical properties of crystals. There are two standard types of dislocation, the 'edge' and the 'screw' dislocation, with a series of intermediate dislocations which can be regarded as mixtures of these standard types in various proportions. It is the screw dislocation, or at least dislocations with a screw component, which are important for crystal growth.

A screw dislocation may be thought of as arising from the following operation on a perfect crystal. A cut is made, as in Fig. 4 (*left*), and the material to the right of the cut pushed down by one molecular diameter. The result is shown in Fig. 4 (*right*): instead of pushing down the material on the right, we might push down the material on the left. In the first case we should have a right-hand screw and in the second case a left-hand screw dislocation. Both cases give a step in the surface of the crystal which extends from an edge to a point *a* in the surface where the axis *ba* meets the surface. At a distance of a few intermolecular distances from the axis, the crystal appears to deviate very little from perfection: only very near the axis does anything seem to be wrong. In fact, the molecules in the crystal now no longer lie on planes but on a screw surface about the axis. This can be seen by considering a route, taken from the top of the step, going round the point *a* back to the bottom of the step. We have followed a path on the surface, which used to be flat, and instead of finishing at the starting point, we finish one molecule beneath it. If the route is continued, so that the layer of molecules below the first layer comes into consideration, we shall end two molecules

below the starting point. The route round the axis *ab* is like a spiral staircase.

If the vapour surrounding such a dislocated crystal is above the equilibrium value growth takes place by the forward movement of the step as before, but now the step does not disappear, but simply rotates like the hand of a clock about the axis

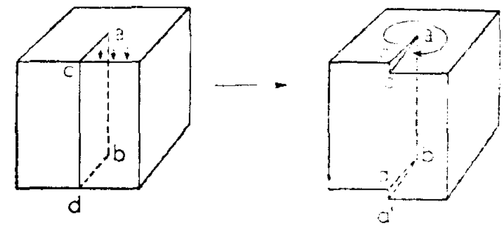


Fig. 4. Illustrating the formation of a screw dislocation with axis *ab*. Permanent steps *acc*^{1}, *bdd*^{2} are left in the upper and lower surface respectively

(actually, in Fig. 4 the rotation would be anticlockwise: it will be clockwise for a left-hand dislocation). After one revolution, a complete new layer of molecules will have been added, but the step will be in its initial angular position.

In practice, the motion of the step will not really move like the hand of a clock, for every portion of it will tend to move forward at the same rate. The end *a* of the step remains fixed, so the successive positions of the step will be as shown in Fig. 5. That is, the step will be wound into a spiral during growth. The growing spiral will appear to rotate.

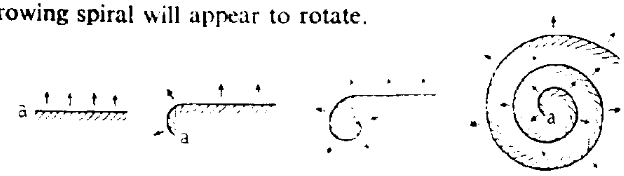


Fig. 5. The growth of a step attached at one end to a dislocation, *a*. Since each portion of the step tends to move forward at the same rate, the step eventually acquires a spiral formation. The shaded side of the line is one molecule higher than the unshaded side: *cf.* Inset, 21–24

If a crystal surface contains a spiral step like that in Fig. 5 the surface will not be flat, but will have a low cone or pyramid on it. Such features are a commonplace on mineral crystals.

It can be shown that the higher the supersaturation the more tightly wound the spiral becomes, and at a supersaturation of about 1 per cent or even less, the distance between the turns of the spiral becomes less than the mean migration distance of surface-adsorbed molecules. Under these conditions the crystal grows as if the surface were entirely covered by kinks, that is at the maximum rate. If the supersaturation is reduced, the spiral unwinds, and the distance between turns becomes greater than the mean migration distance, and the growth rate becomes less than the maximum. Thus the results of Volmer and Schultze's experiments are completely explained, if the crystal contains even one screw dislocation. What if there are more dislocations? The answer is, that unless the number of dislocations is enormous, the behaviour is practically the same as for one dislocation. For illustration consider the case of two dislocations of opposite handedness (Fig. 6). They will be connected by a step, shown in plan in the right hand part of the figure. The growth of such a step takes place by the process shown in Fig. 7. At *c* the turns from the two spirals meet, and where they meet there is no difference of level, so that a closed step separates off, leaving the

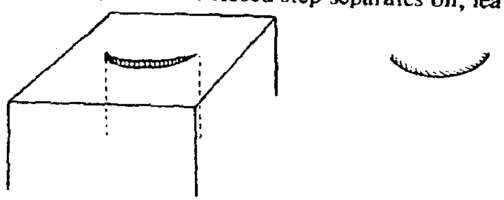


Fig. 6. Two views of a step connecting two screw dislocations of opposite sense

situation shown in *d*. The process continues, more and more closed loops being emitted. This again gives rise to a cone or pyramid on the crystal surface.

The considerations of this article were put forward to a conference held at the University of Bristol in 1949. Our surprise

was considerable when Dr L. J. Griffin, then at the Royal Holloway College, London, dramatically presented to the conference some photographs, obtained through an ingenious optical technique, showing crystals of beryl whose surfaces contained features like those of Figs. 5 and 7. Two of his photographs are reproduced as Inset, 21–22. He was able to show that

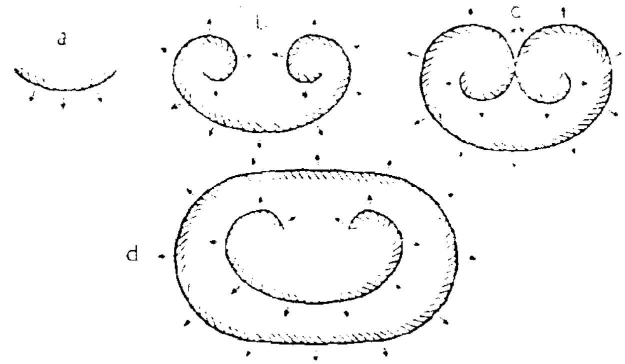


Fig. 7. The growth of a step connecting two dislocations of opposite sense, showing the mechanism whereby closed loops of step are emitted from the dislocation-pair: *cf.* Inset, 21

the lines in his pictures were the edges of steps, and the fact that their ends could be seen terminating in points establishes beyond reasonable doubt that the Frank picture must be substantially correct. More lately, Dr I. M. Dawson and Dr V. Vand of the Chemistry Department at Glasgow University have published an electron-micrograph of a paraffin crystal, grown from solution in the laboratory, which shows the same spiral-type markings. The spiral makes six complete turns from centre to edge, and is of almost perfect regularity (Inset 23). Other micrographs are described as showing the effects of multiple dislocations on the growth of the crystals (Inset 24).

Of course, we have not established that all crystals grow in this way, but only that a reasonable theory, which is in agreement with a large body of facts, can be built up on the assumption that crystals do grow in this way, and that a perfect crystal could not grow at all at low supersaturations. Griffin's evidence shows that some crystals, at least, make use of the dislocation mechanism. No other satisfactory mechanism has been proposed, as far as I know.

The theory of growth from a melt is not covered by the above considerations, and further work is necessary before an overall picture can be given.

REFERENCES

Volmer and Schultze. *Z. Phys. Chem.* (A) 156, 1 (1931)

Burton, Cabrera, and Frank. 'A Theory of Growth of Real Crystals.' *Phil. Trans.* 'A' 243, 299–358 (1951)

RADIO ASTRONOMY

A. J. HIGGS

It is still a popularly-held belief that *millions* of stars may be seen on a clear night. This of course is erroneous: the unaided eye can distinguish only 5,000 or so separate point sources of light in the night sky, even under the best seeing conditions. It takes the combined efforts of powerful telescopes and sensitive photographic plates to establish that the number of separately distinguishable stars runs into millions. We know a surprising amount about these other worlds around us, much of this based on rather slender clues, namely, from a study of that small sample of the energy they emit which is centred around the special region of wavelength to which our eyes happen to be sensitive. In the past few years, however, a new astronomical tool has been forged for us, and a new branch of science has arisen. It has come to be called Radio Astronomy: it uses radio aerials in place of conventional telescopes and radio receivers at their foci instead of eyepieces.

The elementary particles of which all matter is made up are in a state of constant movement and, as a result of these movements, they emit 'radiation' in the form of electromagnetic waves. We are quite familiar with this radiation when it is of the right wavelengths to produce the sensations of light and heat, as from an object which is hot enough to shine brightly. But that atoms in motion also emit, at the same time, waves which can be picked up on a radio receiver is a much less familiar idea. The signal strengths are so low compared with those from radio stations that until recently it was quite impossible to hear them, and so it is not surprising that hitherto they have gone unnoticed. As a result, however, of the very great advances in radio techniques which took place during World War II, we can now build radio receivers which are sensitive enough to detect and measure the radio frequency radiation emitted by all bodies; provided, of course, that it is not swamped by stronger signals. The earth is constantly bathed in the relatively weak but quite unmistakable radio waves which reach us from the sun and from the depths of space. By tuning-in to the creaks and groans of the universe around us, we can obtain quite a new source of information about the stars: how they are born, live and die.

METHODS OF RADIO ASTRONOMY

There are two basic methods of observing in radio astronomy. In the first, we use the familiar processes of radar to explore space around us by sending out powerful signals and looking carefully at the echoes when they come back. From its nature this method is limited to nearer distances and has so far been used successfully only on meteors and the moon. It would probably provide extremely interesting information if we could apply it to the sun, but the technical difficulties are considerable, and the experiment has not yet been tried. The echo method is limited, in any case, to the solar system.

The second method is the one in general use. Radiation of radio wavelengths, emitted in the normal course by extra-terrestrial bodies, is collected by a directive or 'beamed' aerial system (so that we can find the direction from which the waves are coming), and is passed to a sensitive receiver with means for measuring their strength and recording their variations. The range of wavelengths used is from about 1 centimetre to 10 metres. Longer wavelengths are prevented from reaching the earth by the ionosphere and shorter ones by absorption in the gases of the earth's atmosphere. It is usual to mount the aerial system so that its beam, or 'cone of reception', may be pointed in various directions, and often a clockwork or motor drive is fitted to counteract the earth's rotation and thus keep the aerial pointing at the same direction in space. This combination of aerial-plus-receiver-plus-recorder is often referred to as a radio-telescope. The analogy with an optical telescope is quite close, as may be seen from Fig. 1, or from Inset 25 which shows a typical system.

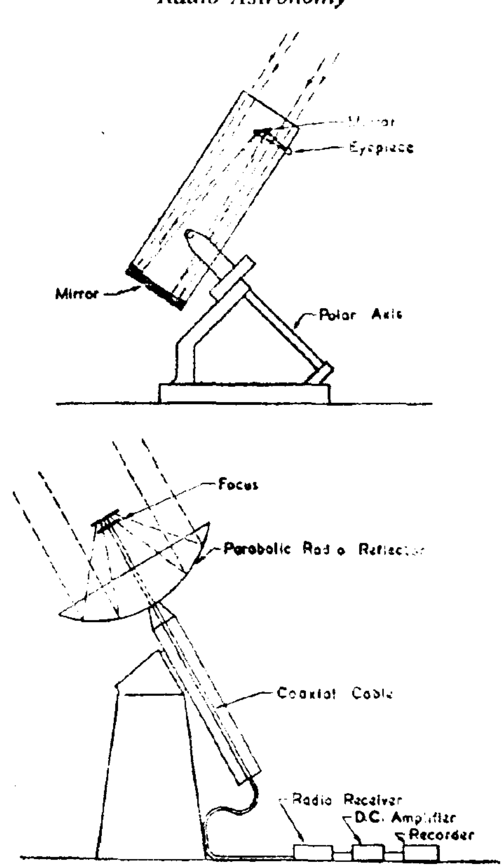


Fig. 1. Diagrammatic representations of reflector telescope (above) and radio telescope (below)

Because it is not feasible to make radio-telescopes with an aperture-to-wavelength ratio which even approaches that of an optical telescope they 'see' an area of the sky which is measured in degrees or tens of degrees across. Measurements of angular position are therefore much less precise; on the other hand, the power sensitivity is quite remarkable. It is possible, for example, to detect radio waves from a body at a temperature only a few degrees above absolute zero.

INTERPRETING THE RESULTS

The information provided by radio-telescopes is of two kinds. First, they tell us the direction, and in some cases the position in space, from which the waves are coming. A relatively rough indication of direction is obtained by swinging the aerial until it gives its strongest signal, but unless the aerial system is larger than convenient the cone of reception may be up to several degrees wide. More precise data can be got by interference methods, for example by combining the signals received by two or more radio telescopes separated by an accurately known distance. In this way the position of the stronger sources of extra-terrestrial radio waves can be fixed to a small fraction of a degree in the two co-ordinates usually used by astronomers, declination and right ascension.* Radio telescopes, then, confirm that radio waves *do* reach the earth from outside, and from which directions. But they do much more than this. The second kind of information they provide is derived from measurements of the power behind these waves. To enable us to understand how this is possible, we must digress for a moment.

We have already mentioned that the natural vibrations within atoms and molecules give rise to the emission of electromagnetic radiation. Some idea of the very wide frequency range covered is given by Fig. 2*a*, in which only the narrow section marked 'visible' can be seen or felt by our unaided senses. Now the laws governing these emissions are well known for what we call 'black' bodies, that is bodies which are opaque to and absorb without reflexion any radiation which falls upon them. The

* Just as we can specify a point on the earth's surface by its latitude and longitude, so the position of a star can be specified if we regard the heavens as a 'celestial' sphere with the earth at its centre. Declination is the equivalent of latitude, and Right Ascension of longitude, and the celestial poles are the points where a line joining the North and South poles of the Earth cuts the celestial sphere.

amount of energy they give out at any part of this range of wavelengths depends on their temperature. Fig. 2*b* shows how this varies for two typical bodies, one at the surface temperature of the sun (6,000°C) and the other at the surface temperature of the earth (30°C). (Note that logarithmic scales have been used to compress the size of this figure. The ratio of the longest to the shortest wavelength shown is ten thousand million (10^{10}), and of the power scale 10^{25}.) By measuring the energy received at any point in the electro-magnetic spectrum we can immediately form an idea of what the temperature must be at the place where it originates. Because temperatures arrived at in this way are not directly measured, but are inferred on the assumption that the source of radiation is a 'black' body, we usually refer to them as 'equivalent' or 'apparent' temperatures. All emitters of radiation are not, of course, black bodies in the above sense. Some gaseous regions, for example, are not completely opaque: we know that these, while absorbing some of the radiation which reaches them, can make a characteristic contribution of their own the intensity of which depends on their 'opacity' as well as on their temperature. The important point is that if the amount of radiation can be measured which is received at various wavelengths, much can be discovered about the conditions which gave rise to it.

The radio waves used in radio astronomy are not different in nature from light waves, but they differ enormously in both wavelength and behaviour. This is why radio methods have provided something of a revolution in astronomy. It is well known that infra-red rays penetrate haze much better than do light waves. Radio waves show the same tendency but in much greater degree; with them we can not only 'see' easily through cloud (and so can observe in any weather), but also through huge patches of dust which are present in interstellar space. To light waves they are impenetrable, as is evident from apparently coal-black patches along the Milky Way. Conversely, radio waves can be reflected or absorbed by ionized gases which may be transparent to light. The earth's ionosphere, the ionized regions high in the atmosphere, is an example near to hand. It is quite invisible,

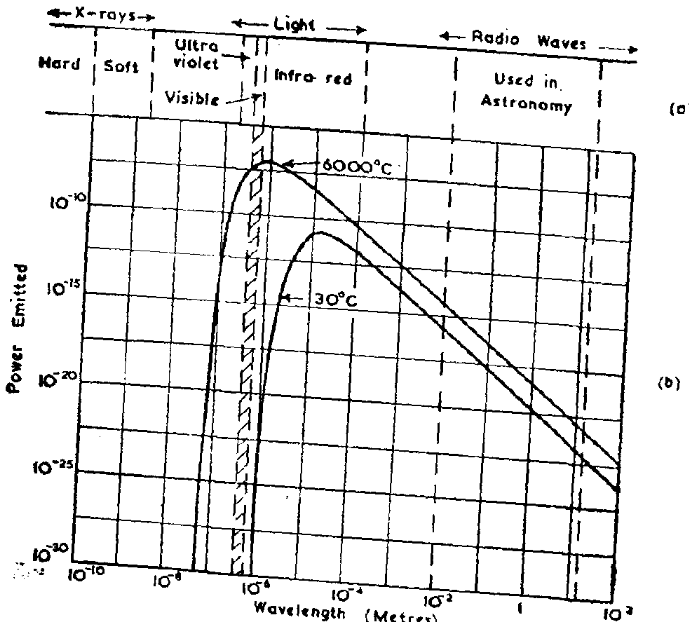


Fig. 2 (*a*). The electromagnetic spectrum, showing approximately the wave-length ranges of *x*-rays, light and radio waves
(*b*). How the power emitted varies with wave-length for two 'black' bodies at temperatures of 6,000°C. and 30°C. (the power is expressed in watts per square metre of surface area per unit frequency-range of the spectrum)

yet responsible for all radio communication over long distances. This property turns out to be specially valuable, because the outer atmospheres of the sun and the stars are extensively ionized.

RADIO WAVES FROM THE MILKY WAY

It was K. G. Jansky, an engineer of the Bell Telephone Laboratories in the U.S.A., who first discovered that radio waves could be detected from sources outside the earth. This was in 1932.

He was tracking down sources of interference to short wave radio communication circuits and noted that a steady hiss seemed to come from the direction of the sun. Observations spread over most of a year showed that the sun was not responsible, but that the radiation came from the general direction of the Milky Way: the sun had just happened to be in that part of the sky. He inferred that 'the source of these radiations is located in the stars themselves, or in the interstellar matter distributed throughout the Milky Way.' Because these signals produce the same sort of sound in a loud speaker as the valve and circuit 'noises' with which radio engineers are familiar we refer to them generally as *cosmic noise;* as *galactic noise* if we are speaking of those which originate in our own galaxy, or as *solar noise* for the similar radiation which, as we shall see later, comes from the sun.

These more or less accidental observations by Jansky opened up a new era in astronomical investigations, but it was not until 1940 that a deliberate attempt was made (by G. Reber, in the U.S.A.) to study cosmic noise with apparatus specially designed for the purpose. The war left us with a fine heritage of highly sensitive receivers and directive aerials, some of them almost ready-built as radio telescopes, and an invaluable fund of engineering knowledge on metre and centimetre wavelengths. Progress has been very rapid in the last few years, and the way in which the intensity of cosmic noise varies in different directions in space is now well known.

The distribution of radio noise over the sky follows pretty closely what is known of the distribution of stars in the extensive star system to which we belong – the Galaxy. The exact shape of the Galaxy cannot be obtained directly from ordinary astronomical observations, because our view in some directions is obstructed by dust clouds. Since these are much more transparent to radio waves than to light, radio methods look like providing a truer picture. There are good reasons, for example, for believing that the Galaxy shows the same sort of spiral structure as is characteristic of many of the very distant external nebulae or 'island universes' (*See* Inset 26). Visual observations tend to support this belief, but the radio observations make it highly probable. They show the sun to be situated in or near one of the spiral arms, and suggest strongly that the direction of rotation of nebulae is that of a spiral unwinding. It seems probable also that cosmic radio noise originates not only in the stars but in the clouds of hot gases which exist between them.

RADIO STARS

Fig. 3 shows part of a typical record taken with a radio-telescope. What the aerial picks up is amplified in the receiver and made to operate a pen whose deflexion upwards from the bottom of the record depends on the strength of the signals. The pen presses against paper which is pulled forward at a constant speed, so that the curve traced shows the variation with time in the strength of the noise-signals. Time marks are put on the record, because accurate timing plays an essential part in determining the origins of these radio waves. The particular record of Fig. 3 was made from a high cliff overlooking uninterrupted sea near Sydney, Australia, with the telescope pointed at the eastern horizon and fixed in position. As the earth rotates, different portions of the sky 'rise' above the horizon and pass through the beam of the radio-telescope, that is through the narrow cone over which it responds.

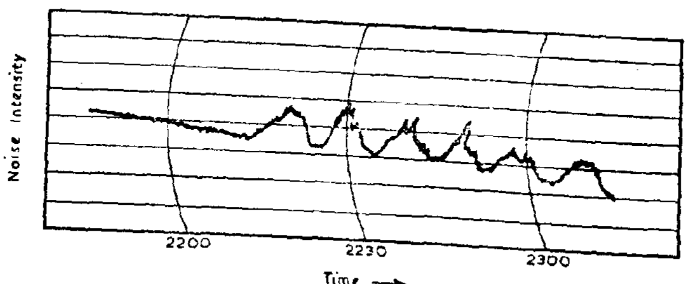


Fig. 3. Portion of a cosmic-noise record made with a radio telescope pointing in a fixed direction. The generally downward slope from left to right is due to a bright portion of the Milky Way gradually passing out of view, and the series of humps and troughs commencing suddenly about 2230 to the rising of a radio star at that time

Notice, first, that the trace slopes generally downward from left to right: a bright part of the Milky Way was passing from 'view' as the record was taken. But notice, now, the regular series of humps which begin suddenly at about 22.15 hours: what do these imply? Remember that this telescope is specially placed on a high cliff overlooking the sea. Radio waves from the sky just above the horizon can enter the aerial direct, and they can also do so after being reflected from the flat surface of the sea, having travelled just a little farther in the latter case. The receiver cannot distinguish between direct and reflected waves, and records their algebraic sum: a maximum if they arrive exactly in step, and a minimum when they are out of step. When the origin of the radio waves is moving (rising above the horizon) the direct and reflected waves will be alternately in-step and out-of-step, but this will be noticeable only if the source is spread over a very small region of the sky. A record such as that of Fig. 3 therefore indicates quite clearly that, through a background of more general cosmic noise, a source of very small area appeared on the horizon at about 22.15 hours. If in the same way we find the time of its setting below the western horizon, we can determine its exact position in the sky – its Right Ascension will be the sidereal time when it was overhead (half way between rising and setting) and its declination can be found from the time taken to cross the sky. The sea provides a convenient means for accurate position-finding because only one aerial is required. The same sort of interference pattern can be obtained from the use of two or more aerials spaced a known distance apart and combined in pairs.

Upwards of fifty point sources have been discovered by interference methods. They have been christened 'radio stars' because in most cases their positions do not coincide with those of any known heavenly bodies. There are, however, several important exceptions. The well-known Crab Nebula (Inset 27), believed to be the expanding shell of an exploded star first observed as a supernova by Chinese astronomers in 1054 A.D., coincides closely in direction with that of a source in the constellation Taurus. One of two known sources in the Centaurus region has been provisionally identified with another somewhat unusual astronomical object, the extra-galactic nebula NGC 5128 (Inset 28); while English observers have just reported success in detecting radio noise from that magnificent spectacle of the skies, the Andromeda Nebula. Radio-stars are *not* the optically bright ones which may be seen on any clear night. The Crab Nebula is at a distance from us of something like 4,200 light-years. This is enormously further than the sun, which is at a distance of only 8 light-minutes, yet the radiation we receive from the Crab Nebula on a wavelength of 3 metres is about one tenth of that from the sun under quiet conditions. Another powerful source in Cygnus presented, for some time, a puzzle of another kind; it appeared to fluctuate in intensity. This twinkling effect, however, turns out almost certainly to be produced in our own atmosphere, much as ordinary stars are seen to twinkle. The record reproduced in Fig. 3 is of Cygnus rising, and the variations in intensity can be seen on the crests.

It is not known how these high powers are generated but, in the case of an exploding star at least, it is easy to imagine that there may be plenty of energy available for conversion into radio waves. Despite their great power, however, the known radio-stars provide only a small fraction of the total cosmic noise which reaches the earth, and without the refinements of the interference methods we should have been unable to separate their contribution. Probably many fainter ones still remain to be discovered, and it is possible that radio-stars may be responsible for a substantial part of all cosmic noise, most of them being too distant to be distinguished.

RADIO WAVES FROM THE SUN: SOLAR NOISE

The sun forces itself upon our attention as something special in the heavens only because it is so close, relatively. The amount of light and heat it provides is immensely greater than that from all the other stars put together, yet astronomers know that, as a star, it is below average. As an emitter of radio waves it is quite insignificant compared with the radio-stars of which we have just been speaking.

The discovery of cosmic noise (in 1932) prompted attempts to detect similar radiation from the sun. But radio was not then sufficiently advanced, and it was not until 1942 that radio waves from the sun were noticed – again, more or less accidentally. This was in England during the recent war when radar sets were constantly scanning the sky for enemy aircraft. For a few days in February of that year they picked up peculiar signals quite unlike the echoes usually observed from aircraft. Their directions pointed consistently towards the sun, on which an unusually large sunspot was then visible. Since time immemorial, spots on the sun have often (but usually unfairly) been accused of causing peculiar happenings on the earth. This time, as we shall see later, the blame was laid correctly. Later in the same year, some of the new 3-centimetre and 10-centimetre radar sets, built for military use, were turned on the sun by an observer in America. He succeeded in finding radio waves which without doubt were coming from the sun, but were steady from day to day. These two separate observations of radiation from the sun in the radio spectrum, one relatively steady, and the other varying widely in association with sunspots, aroused very great interest. In the intervening years a good deal of careful work has been devoted to radio-studies of the sun, particularly in Australia.

THE QUIET SUN

The sun is surrounded by gases which become more and more tenuous with increasing distance. It therefore has an atmosphere, but one very different from that of the earth, because at a much higher temperature: so hot, in fact, that neutral atoms and molecules are almost completely broken up into electrons and ions. It is usually divided into three fairly well defined regions. The photosphere is what may be seen with the naked eye and defines the sun's visible disc: it is at a temperature of about 6,000°. Above this is the chromosphere which extends out a further 6,000 miles or so and appears to be at a temperature of about 30,000°. Spectacular upsurgings of gases in the chromosphere, known as prominences, are sometimes seen: these may shoot out to distances of the order of a few hundred thousand miles. The outer regions, reaching to a distance of something like a million miles from the sun's surface, are known as the corona. Its temperature, before radio astronomy, was thought to be not higher than that of the chromosphere, but now we know that it is very hot indeed; with a temperature of a million degrees in parts.

Those fortunate enough to have seen some of the striking moving pictures which have been made of the sun's atmosphere, cannot fail to have noticed that it is in a state of almost continuous activity. It may seem something of a misnomer, then, to refer to the sun as 'quiet'. Fig. 4 shows two examples of records made of solar noise, the upper recorded at a wavelength of $1\frac{1}{2}$ metres and the lower at 1 cm, the sharp dips at *A* and *B* being caused by momentarily pointing the aerial away from the sun. Notice how steady is the intensity of the 1 centimetre solar noise and how variable is that at $1\frac{1}{2}$ metres; but that between the wild fluctuations the recording pen does not come back to zero, it is still recording something. We regard this *base* level which hardly ever varies at the shortest wave-lengths and, at the longer wave-lengths, is always present between fluctuations (sometimes without them), as the radiation from the quiet or undisturbed sun. We have seen that anybody radiates electromagnetic energy over a very wide range of wave-lengths because of the 'thermal' motions of its atoms, and at an intensity depending on its temperature. This base level which we pick up from the sun with our improved modern receivers is simply the sun's 'thermal' radio energy. We can use our knowledge of the laws of radiation to find the equivalent temperature of the region from which it comes. Neither the older optical nor the newer radio method is able alone to give us a picture in depth of the sun's atmosphere, but together they do so. The coronal regions are highly ionized and, like our own ionosphere, completely prevent radio waves longer than a certain wavelength from passing through them. This point of cut-off depends on the number of electrons present. Optical measurements tell us the electron density at various heights and so we can find out where the radio waves of different wavelength which escape from the sun must have originated. The longest wavelengths can come only from the outermost regions of the corona,

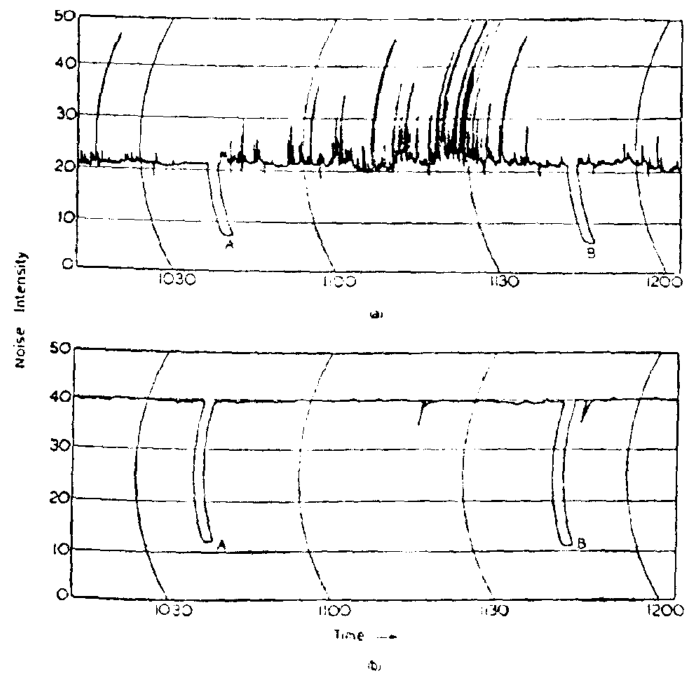


Fig. 4. Two typical records of solar noise. The upper is taken at a wave-length of 1·5 metres and shows the rapid variations which are usually characteristic of the longer-wave radio radiation from the sun. The lower record was made at a wave-length of 1 centimetre; at these very short wave-lengths sudden fluctuations hardly ever occur. The sharp dips at *A* and *B* were produced by momentarily pointing the aerial away from the sun: they show that what the radio telescope was recording came from the direction of the sun

the shortest from as low as the chromosphere. As a result therefore of this happy combination of radio and optical data we are now provided with a picture of how the temperature and density of the sun's atmosphere vary from top to bottom. The fact that parts of the corona attain the astonishingly high temperature of a million degrees is thus confirmed. Observations have not been continued long enough to show whether the sun as a whole is heating up or cooling down, but Australian observers are finding that *parts* of the corona are not as hot now as they were a few years ago. This effect, probably, has something to do with the sunspot cycle. A prediction which is in course of being verified is that at some wavelengths the rim of the sun should appear hotter than the centre – so that if we had 'radio' eyes we would see the sun not as a uniform disc, but as a bright circle with a less luminous centre.

SUNSPOTS: THE DISTURBED SUN

The solar noise record reproduced in Fig. 4*a* makes it clear that, on that day at least, the sun was emitting additional radiation on a wavelength of 1½ metres over and above that arising from thermal causes. Records taken on other days, over long periods, and on other wavelengths confirm that this is usually, but not always, the case. At the shorter wavelengths (Fig. 4*b*), it is remarkably steady. What the English radar observers noted in those dark days of February, 1942, gives a clue to the situation – sunspots. This was confirmed by reports from short-wave listeners who over some years around sunspot maximum had noticed that reception was sometimes marred by the presence of a background of rushing or hissing noises. Careful series of measurements have since established that the average intensity of solar noise is likely to be greater when sunspots are visible, and that this extra, or 'enhanced', radiation comes usually from small areas near them. The correlation is not always exact, however: in the few years over which the sun has been closely watched giant sunspots have *always* been associated with more intense radiation, but many smaller ones have not.

How can we be sure of this apparent connexion between enhanced radiation and sunspots, and why is it much more noticeable on long-wave rather than short-wave radiation from the sun? In Fig. 5 the intensity of solar noise (this time at a wavelength of 25 cm.) has been plotted day by day over a three-month period, and below it the total area of the sunspots visible on each day. The two curves follow each other very closely, suggesting that the

two phenomena are correlated in some way. Two methods have been used to see if the additional radiation when spots are numerous comes from the disc as a whole or from localized regions. The first can be applied easily when there is only one major active area on the sun, as is often the case at the longer wavelengths. It uses the more precise interference methods described earlier and shows quite definitely that enhanced radiation *does* originate in small areas, and that these are usually near sunspots, or where sunspots have been seen before. The second, though ingenious, can be used only when the sun is in eclipse: it is more applicable at the shorter wavelengths, where fluctuations are usually absent.

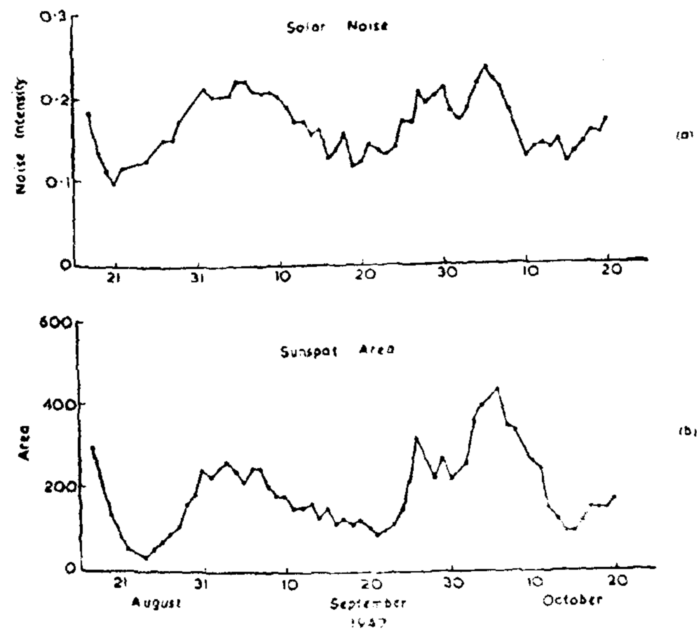


Fig. 5. Correlation between enhanced solar noise and sunspots. The upper curve shows how the daily average of solar noise, at a wavelength of 25 centimetres, varied over a 3-month period. Below are the corresponding variations in the area of visible sunspots

As the moon passes across the sun it will cover up one by one any specially 'active' areas, and the solar noise will decrease suddenly. If we are observing at one place all we can say is that each of these sources must be somewhere on the sun marked out by the position of the edge of the moon at the time. If, however, we are observing at two or more places some hundreds of miles apart, the moon's apparent position at each of these will be different, the arcs defining them will intersect, and where they intersect is the location of the area we are looking for. Fig. 6 shows the sources which were located in this way during the solar eclipse of 1st November, 1948, in south-eastern Australia. Most of them are seen to be close to the position of visible sunspot groups or where sunspots appeared a month before. Notice particularly, however, that two of them were well off the western limb of the sun.

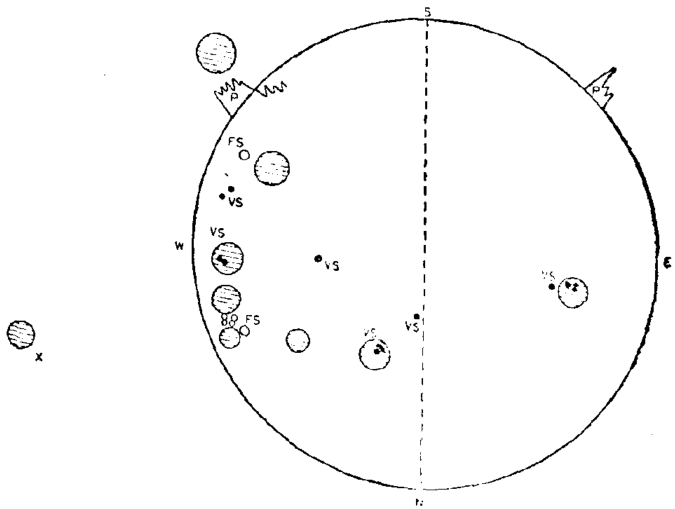


Fig. 6. The sources of solar noise located during an eclipse of the sun which occurred in south-eastern Australia on 1 November, 1948. Most are seen to be close to visible sunspot groups (the black dots, marked *VS*), or to places where spots appeared on the last revolution of the sun (the open circles marked *FS*). Two sources, however, were well off the western limb

It is well known that sunspots come and go with a fairly regular period of 11½ years' duration; that short-wave radio reception varies over the same period; that such happenings as magnetic storms and auroral displays are closely connected with the central passage of large sunspot groups, and so on. Sunspots, in fact, are visible manifestations of far-reaching disturbances in the solar atmosphere. We know from optical observations that very large electric and magnetic 'fields' are associated with them: they may be somewhat analogous, therefore, to thunderstorms, but on a more magnificent scale. The existence of giant magnetic fields in the vicinity of sunspots gives a further proof that enhanced radiation comes from sunspot areas also. Radio waves from a region of pronounced magnetic field are, as we say, 'polarized' in a certain way. Enhanced radiation always exhibits this characteristic polarization. We know, too, that it is generated mainly high up in the sun's atmosphere because it is much more noticeable at longer than at shorter wavelengths; the latter, as we have seen, coming from deep down in the corona. Present ideas suggest that enhanced radiation must be caused by something similar to a lightning flash, but more terrific: if it were thermal in origin, like the radiation from the quiet sun, improbably high temperatures of ten thousand million or sometimes even ten million million degrees would be needed to account for it.

BURSTS AND OUTBURSTS

The level of the enhanced radiation associated with sunspot activity may vary over a wide range (100 to 1 or more) but it does this over periods of hours or days. Such periods are sometimes referred to as 'noise storms'. In marked contrast to these slow changes there are other disturbances which have an extremely sudden commencement and last for only a very short time, from fractions of a second up to perhaps 20 seconds. They have been aptly named 'bursts'. They appear almost exclusively on the longer wavelengths and usually in great numbers during noise storms, but may occur even when the sun appears quiet. And there is another phenomenon which has the same general appearance on our records but is much more intense and lasts longer, minutes instead of seconds: these we call 'outbursts'. Although outbursts look rather like an outsize in bursts they are really quite different.

During the past year important work has been done in Australia with an instrument called a spectrum-analyser which allows these short-period fluctuations to be 'seen' simultaneously over a wide band of wavelengths. Outbursts are shown to be in a class of their own. They occur at the shorter wavelengths first and progressively up to minutes later at the longer wavelengths, always in this sequence. We can interpret this pretty definitely as the effect of an agency moving outwards in the solar atmosphere.

The largest outbursts have always occurred when either a solar flare or a radio fade-out has been observed. A solar flare is the name given to the sudden appearance on the sun of an unusually bright patch which can be seen at wavelengths in the ultra-violet region, generally near a sunspot. We have known for many years that the great increase in the sun's ultra-violet radiation at the time of a flare disturbs the earth's ionosphere and causes a radio fade-out, that is a serious interruption to short-wave communication. We also know that some 24 to 48 hours after a large flare the earth experiences severe magnetic storms, often accompanied by spectacular auroral displays, and that these are due to the arrival in our atmosphere of particles shot out at high speed from the sun at the time of the flare. It is almost certain that the agency responsible for this stream of high-speed particles is identical with that which gives rise to outbursts. Outbursts therefore now join that extremely restricted list of phenomena for which we can associate something observed on the earth with something seen on the sun.

About bursts we cannot be so sure: the spectrum analyser, however, tells us that there are distinct kinds. One of them shows the same characteristic polarization as enhanced radiation and is generally associated with noise storms. 'Storm' bursts are of very short duration and occur over very narrow and stationary bands of wavelengths. They must therefore be due to very momentary disturbances high in the corona, almost certainly of electrical origin. In fact it has been suggested that, much as the separate lightning flashes in a very severe but distant thunderstorm can produce an almost constant flicker of light in the sky, so a solar 'noise storm' may be the result of a very large number of practically simultaneous storm bursts, each arising from the solar counterpart of a lightning flash.

The second kind of burst is equally short-lived and therefore also has its origin in a very localized event in the corona. They may occur at any time. They are not associated with noise storms, but some of them show a very rapid change in wavelength, much greater than that of outbursts. All that can be suggested at the moment is that these 'isolated' bursts may be due to the rapid dispersion of a dense cloud of ionized gas, or to the motion outwards through the solar atmosphere of a disturbing source. If the latter, its speed must be much greater than that of any known movements of the gases of which the corona is composed.

THE MOON

The moon is of less interest than the sun and the stars because it is a 'dead' body. It has long since lost its atmosphere and emits no characteristic 'noise' of its own other than as a thermal body at a surface temperature which varies with the lunar phase: from rather less than that of boiling water at full moon, to 150° C. below freezing at new moon. It is, however, our nearest neighbour in space, and is the only one to which we have so far been able to apply both echo and natural-radiation techniques.

That radio echoes can be obtained from the moon has been demonstrated in the U.S.A., in Hungary and in Australia. The Australian experiments are of interest in that standard short-wave broadcasting transmitters and receivers were used to see whether echo signals from the moon had suffered absorption on their double journey. It is clear from these experiments that, apart from our own ionosphere, there are no major absorbing regions between the earth and the moon; and further, that, given a full-time transmitter, moon-echoes could be used to explore the full depths of the ionosphere in a manner not possible by other means.

The temperature of the moon has been studied for many years by measuring the natural radiation it emits in the infra-red region of the spectrum. Some measurements of its radio-emission at the very short wavelength of 1·25 centimetres have recently been made in Australia, which generally confirm the values obtained but indicate that, whereas the infra-red methods give the temperature of the moon at its surface, the radio methods supply the answer for a foot or two below the surface. They therefore show, as would be expected, a much smaller variation from full moon to new moon. They also suggest that the moon must be covered by an efficient heat-insulating material, such as would be provided by a thin layer of dust or volcanic ash.

THE PLANETS

Like the moon, the planets are 'dead' bodies and, with the possible exception of Venus, incapable of contributing any characteristic radiation of their own, apart from the appropriate thermal radiation corresponding with the temperature at or near their visible surface. Measurement of this would require the use of very powerful radio telescopes, and has so far not been attempted.

Venus, however, has a considerable atmosphere, and one which, like our own, contains plenty of water vapour – so much so, in fact, that it is doubtful whether we can ever see its surface because of the clouds above it. It is not inconceivable, therefore, that thunderstorms may occur on Venus and, if so, it is within the bounds of possibility that we may be able to detect the 'atmospherics' from them. (It will be recalled that one suggested origin for 'bursts' of solar noise is gigantic electrical discharges in the sun's atmosphere.) When the methods of radio astronomy come to be applied to the planets, Venus appears to be the one most likely to provide new and interesting information.

METEORS

Many people know meteors as 'falling stars' and find it hard to believe that a normal meteor is seldom as much as one millimetre in diameter. Both they and ourselves are moving through space at high speeds and when they enter our outer atmosphere the relative motion is often so great that the meteor and the molecules

it hits are raised to such temperatures that they glow visibly. The tail of a falling star is merely the glowing vapour trail it leaves behind. It is an extremely rare occurrence for us to run into a meteor which is large enough to escape volatilization and reach the earth as a solid lump.

The visual study of meteors is made difficult by their transient appearance, and by the fact that they are mostly invisible by day and often obscured by cloud at night. Consequently the introduction of a radio method of detection, which operates with or without cloud and equally well by day or night, offers interesting possibilities for studying them. Here, of course, we can only use the 'echo' method, but as the meteors themselves are so minute in size, we cannot expect much of an echo from the solid particle itself. It has long been known that they are partly responsible for some of the ionization of the earth's outer atmosphere, that is, when they collide with neutral molecules they knock electrons from them. It is the trail of ionization which can be 'seen' by radio methods, just as the accompanying emission of light by the 'excited' molecules may be seen by eye. This new method of keeping a 24-hour watch on meteors is being extensively used in England and Canada. The directions from which they arrive can be determined, and by this means not only is more complete information being obtained about already known 'streams', but new families of meteors are being discovered.

CONCLUSION

In this brief survey of the new field of radio astronomy we have tried to show how radio methods are now being used to explore the universe around us, at depths so distant that news takes thousands of years to reach us, as well as at its nearer fringes, less than a hundred miles above our heads. Radio methods not only provide a way of confirming what we already know or suspect from other means, but they often lead to startling new discoveries because of differences between the behaviour of radio waves and those of light. Clouds of interstellar dust are more transparent and of ionized gases more opaque to radio waves, bodies are 'incandescent' at low instead of only at high temperatures, and electrical disturbances are made 'visible'. Radio astronomy, however, is yet an infant-in-arms among the sciences. Born of excellent stock, it has a promising future, but the full part it is to play in life must be left to the future to disclose.

THE COSMOLOGICAL PROBLEM

G. C. McVITTIE

A WRITER who seeks to present to the general scientific public an account of the theory of the astronomical universe is faced by two alternatives. On the one hand, he may follow the traditional procedure in cosmological speculations and, by making guesses as to the nature of the universe, reach conclusions that have the appearance of certainty but that would be regarded by many astronomers as more than debatable. On the other hand, he may present as best he can the observational data and then show how, through theoretical assumptions each of which may be questioned, certain tentative conclusions may be reached. The second alternative will be adopted in the following pages because it seems to be in better accord with scientific procedure. But the reader must be warned that no final and clear-cut answers to the problems of cosmology will thus be reached.

The astronomical objects which are the concern of cosmology are the extra-galactic nebulae, or simply the 'nebulae', whose constitution was a matter for debate even as late as the early years of this century. Were they small masses of glowing gas relatively near to the sun, or immense stellar systems at great distances? The construction of the 100-inch telescope of the Mount Wilson Observatory after the First World War settled this question, for this instrument revealed that certain of the brighter, and therefore presumably the nearest, nebulae could be resolved into individual stars. Certain types of stars of known intrinsic brightness were soon identified and a comparison of the intrinsic and apparent brightnesses of these stars led to the determination of the distances and sizes of a few nebulae. For example, the Andromeda nebula turned out to be a stellar system distant some 750,000 light-years from the sun; it occupies a highly flattened lenticular volume some 70-80,000 light-years in largest diameter. This nebula is of the spiral type, consisting of a central spherical agglomeration of stars around which wind spiral arms of stars. Spiral nebulae also contain clouds of gas and of dust located principally in the central plane of the system. In contrast with the spiral type there are spheroidal-shaped nebulae which appear to contain no gas or dust-clouds but are formed of stars alone; and also irregular types of various kinds which show no clear volume-structure in their stellar distribution. Each of these systems, it has been calculated, contains a mass equivalent to between 2×10^9 and 2×10^{11} solar masses, the spiral nebulae probably being the more massive systems. The sun itself is a member of one of these systems, called the Galaxy, which contains all the stars visible to the naked eye, the stars of the Milky Way and many more besides. The Galaxy is probably a large spiral with the sun in an eccentric position in the central plane. At any rate, the Galaxy certainly contains many gas- and dust-clouds which effectively obstruct the passage of light from the remoter parts of the system to the sun's neighbourhood. Ironically enough, therefore, we are less well-informed about the structure of the nebula in which we are located than about some of the others.

The astronomical universe is regarded as populated by nebulae, chiefly because the largest telescopes do not reveal any class of objects which cannot be identified as nebulae or as objects associated with them. However, this consideration is not a very strong one when we remember that, in the past, each improvement in instrumental technique has altered our conception of the universe. It is therefore useful to make a distinction between the 'nebular system' and the 'universe'. The nebular system contains all those nebulae some at least of whose characteristics can be measured and studied; the universe is the 'totality of things', whether these can be observed or not. Even if we assume that the universe contains nothing but nebulae, we cannot suppose that the universe and the nebular system are co-extensive because, if it is estimated that the 100-inch telescope can usefully observe some 63 million nebulae, yet there is no indication that the supply is

running out. The more powerful 200-inch telescope will certainly extend the bounds of the nebular system.

THE OBSERVATIONAL DATA

We shall now describe the observational data available for the nebular system, emphasizing those which are of interest for the cosmological problem. Since each nebula will be regarded as a unit, it is not necessary to refer to observations which bear on the internal structure of a nebula. We remind the reader that no useful observations can be made on the nebular system by visual means and that all data are obtained by photography. In the first place, the apparent magnitude, m, of a nebula can be measured, which means that its apparent brightness can be assigned a number on a certain scale. This quantity is not easily obtained and the difficulties in determining it will be mentioned presently. The scale of apparent magnitudes is so adjusted that an object of the first magnitude is 2·512 times as bright as one of the second, and so on. The result of this system is that really bright objects have negative magnitudes, the full moon, for instance, being of apparent magnitude −12·5. At the other extreme, the 100-inch telescope can be used to measure the apparent magnitudes of nebulae as faint as +21·03. The second kind of measurement which can be made on the nebulae is a very remarkable one. When the light emitted by a nebula is passed through a prism, the usual spectrum crossed by dark lines is observed. On the average the nebular spectra are similar to the solar spectrum which, incidentally, shows that stars of solar type are very common in the nebulae. But a curious fact also emerges: the dark lines of a nebular spectrum are displaced towards the red (or long wave-length) end of the spectrum as compared with the corresponding lines in the spectrum of the sun. If we denote the amount of this displacement by S, it is found that, on the average, S increases with m, so that the fainter the nebula appears to be, the larger the value of S in its spectrum. The set of numerical data, consisting of corresponding values of S and m, may be called the displacement-magnitude data, and they have been obtained by measurements on some 500 nebulae. This may seem to be a small proportion out of 63 million, but the measurements are not easy to make and are indeed restricted at present to nebulae brighter than m equal to about 18. The nearer nebulae, such as Andromeda, do not show this effect in any significant manner. The phenomenon is known as the red-shift, and its interpretation as a velocity effect is discussed in the section headed *Velocity-distance Relation.*

The third kind of observation which can be made on the nebular system is to count the total number, N, of nebulae visible over the whole sky with apparent magnitude equal to or less than some fixed value of m. Since, for the fainter nebulae, this number runs into millions it is only possible to count the numbers in certain regions of the sky regarded as typical and then estimate for the remainder of the celestial sphere. Successive counts of this kind have been made for values of m not exceeding seven limiting magnitudes beginning at 12·78 and ending at 21·03. One of the great difficulties encountered in making these counts is, as has already been said, the determination of a satisfactory apparent magnitude for each nebula. If we had perfect instruments, if the earth's atmosphere were perfectly transparent, if there were no obscuring dust-clouds in the Galaxy, it would be a relatively easy matter to say exactly how bright a nebula appeared to us to be. But even when these obstacles have been overcome by making suitable corrections to the apparent magnitudes, there still remains another difficulty introduced by the red-shift. This difficulty arises from the fact that the photographic plate is not equally sensitive to light of all wave-lengths, so that two nebulae with identical spectra, one of which shows a large red-shift and the other none, will not make the same impression on the plate. The theory of this effect has not been exhaustively investigated, but it is believed at present that the relevant correction amounts to subtracting KS from the apparent magnitude, where K is a number approximately equal to 2. We shall see later, when we deal with the distances of the nebulae obtained from formula (1) on page 67, how this correction enters into the calculations. For the moment we note that the set of values of N and m provide us with a

1. Scientific stations are maintained on Heard I. (about 54° S., 74° E.) and on Macquarie I. (54° 30′S., 158° 57′E.) by the Antarctic Division, Australian Dept. of External Affairs. Corinthian Bay, Heard I., starting point for the annual census of sea elephants.

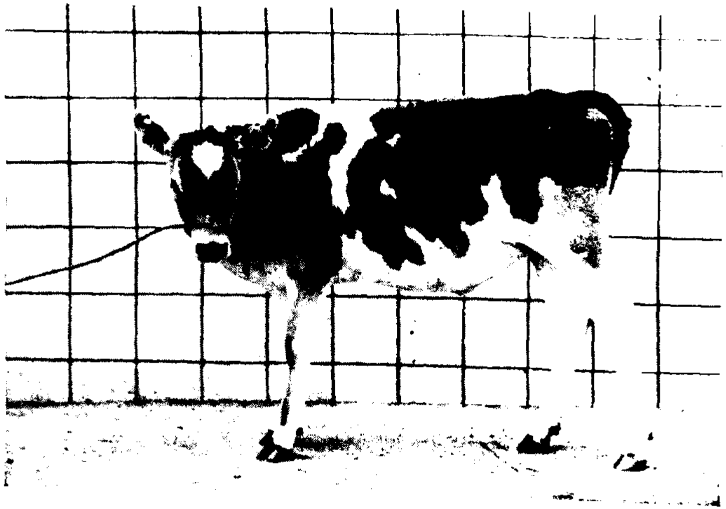
2. Heard I., showing Spit Bay (top left), where the census is taken. (*These and following photographs are reproduced by courtesy of the Australian News and Information Bureau.*)

3. Heard I. Release of daily radio-sonde as part of meteorological programme.

4. Macquarie I. Elephant seals fighting. Here, as on Heard I., a weather station is maintained.

5. Macquarie I. Leopard seal, with teeth well displayed.

6. Macquarie I. Group of elephant seal cows.



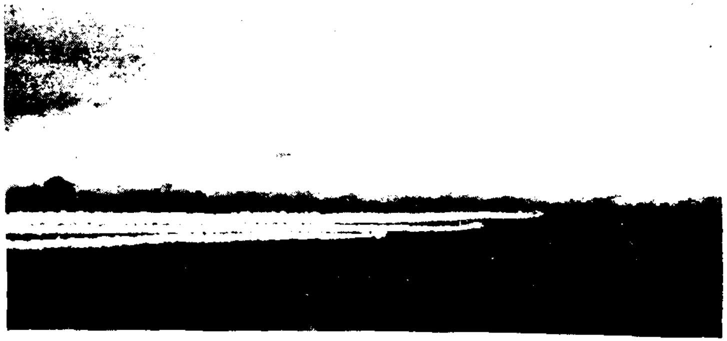
7 and 8. Pair of identical twin cattle (*T*. 67–68 in text tables) at Ruakura Animal Research Station, New Zealand. Note slight variations in white markings.

9 and 10. Further pair of identical twin cattle (*T*. 69–70 in text tables).

11 and 12. Pair of identical twin cattle, showing abnormality of lower jaw. See text for discussion (*T*. 61-62).

13. Main building of South African Veterinary Research Institute at Onderstepoort, near Pretoria.

14. Onderstepoort. The garden of poisonous plants. A demonstration is given of varieties of wild tulip which are widely distributed in South Africa, often mixed with grass. They affect the heart and intestines of cattle which eat them.



15. An uncompleted experiment in the control of tsetse fly. Aircraft applying D.D.T. as a smoke over the Mkuzi Reserve, Zululand. Successive treatments resulted in the reduction of the fly population in the proportion of 9,000 to 3. The aim is the total elimination of the fly within a closed area, south of the main tsetse belt and further results, including effects on other insects, are still awaited.

16. Tsetse fly control. Use of a helicopter to apply D.D.T. as smoke in undulating terrain. *(This and preceding photographs are reproduced by courtesy of the Director of Information, South Africa House.)*

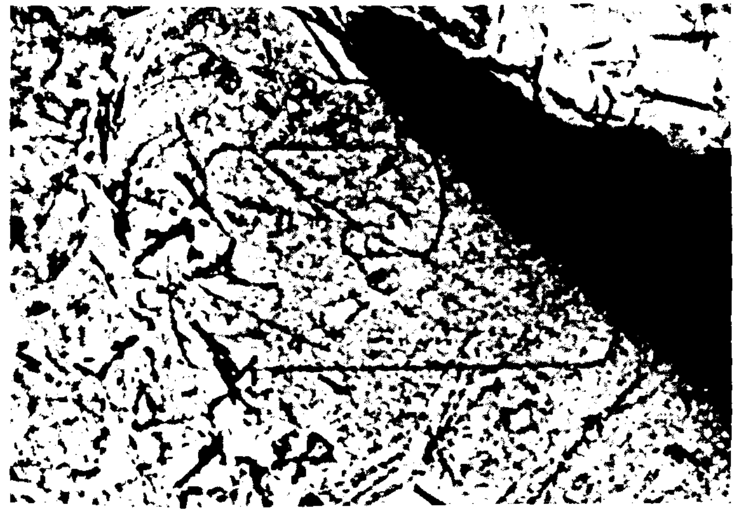
17. Central Glass and Ceramic Research Institute opened last year at Jadavpur, Calcutta. This is one of a series of eleven national laboratories, of which seven have so far been opened. The remaining four are working in temporary accommodation.

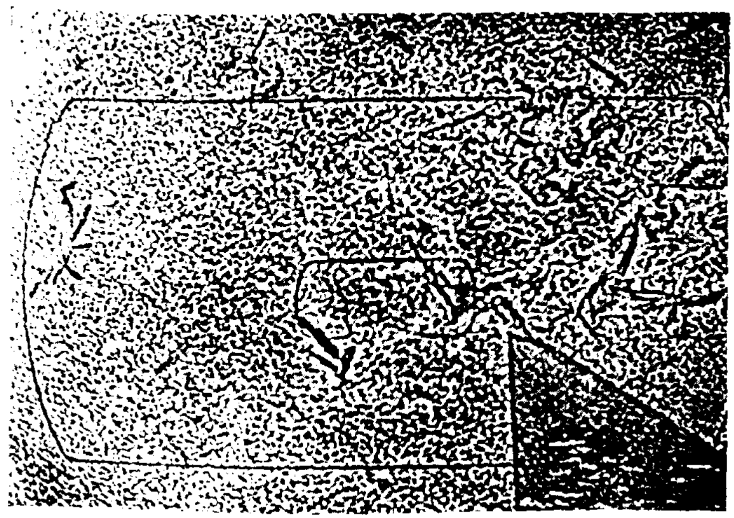
18. Forestry Research Institute, New Forest, Dehra Dun, United Provinces. This is an older laboratory. *(These and the two following photographs are reproduced by courtesy of the High Commissioner for India.)*

19. Ethnographical preparatory room, Department of Anthropology, Benares. As well as tribal institutions, the Department is studying skeletal remains from Harappa, Mohenjodaro and other prehistoric sites.

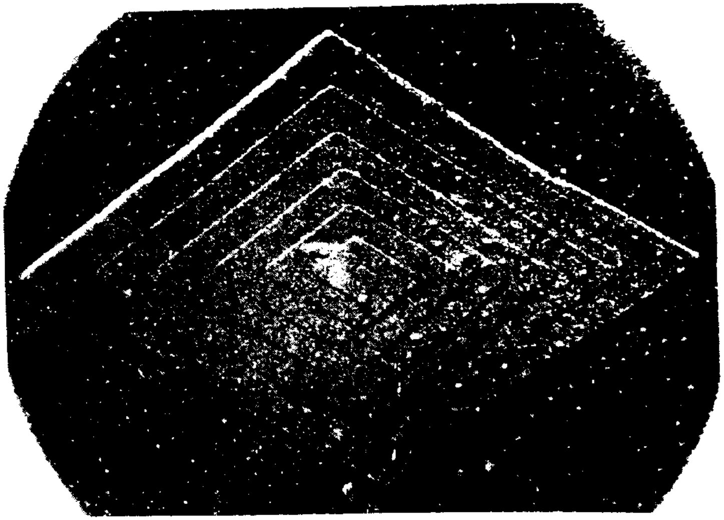
20. Dairy Research Institute, Bangalore. A buffalo and calf of the Nili breed. The raising of milk-yields is a pressing problem: the average milk yield in India is about one-seventh of that in Great Britain.

GROWTH OF CRYSTALS

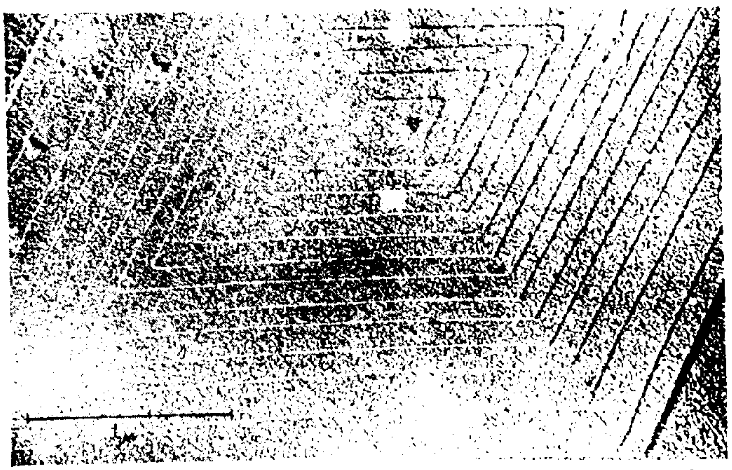




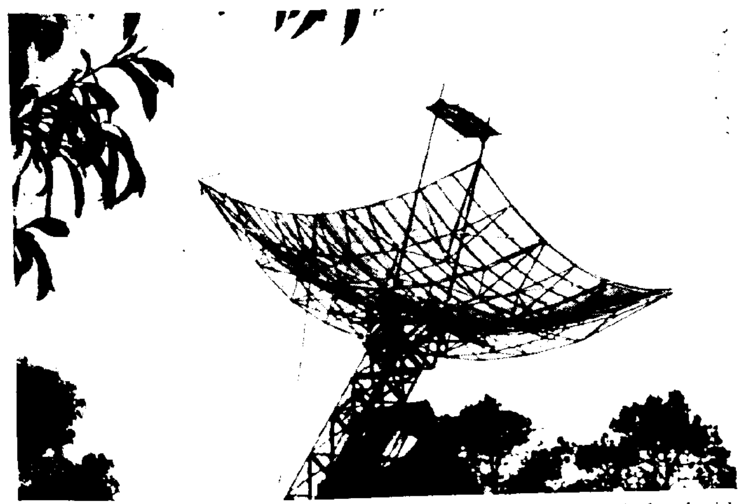
21 and 22. Step markings on the surface of natural beryl crystals, showing mechanism of crystal growth (*cf.* Fig. 5 in text). Magnification × 540 and × 690 respectively. (*Photographs by J. Griffin.*)



23. Electron micrograph of a single crystal of the paraffin, *n*-hexatriacontane, showing the simplest type of spiral growth. Specimen shadow-cast with palladium. Magnification × 20,000. (*Courtesy of the Royal Society.*)

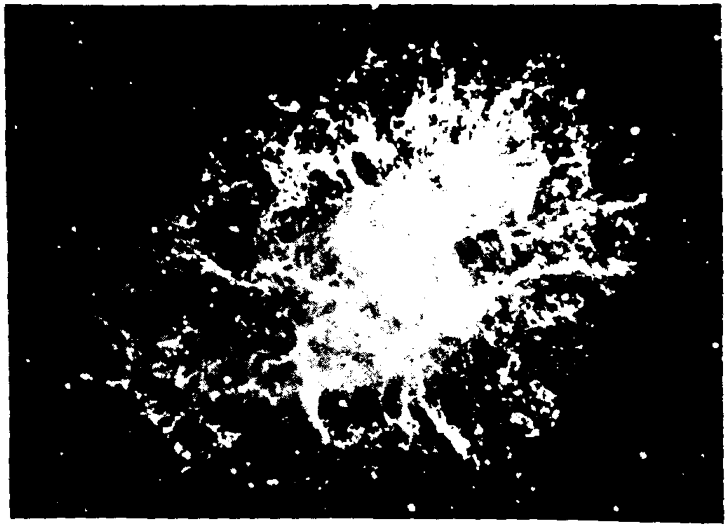


24. Similar electron micrograph to the above, but showing growth from two dislocations of opposite sense in close proximity. Magnification × 27,000. (*Courtesy of the Royal Society.*)

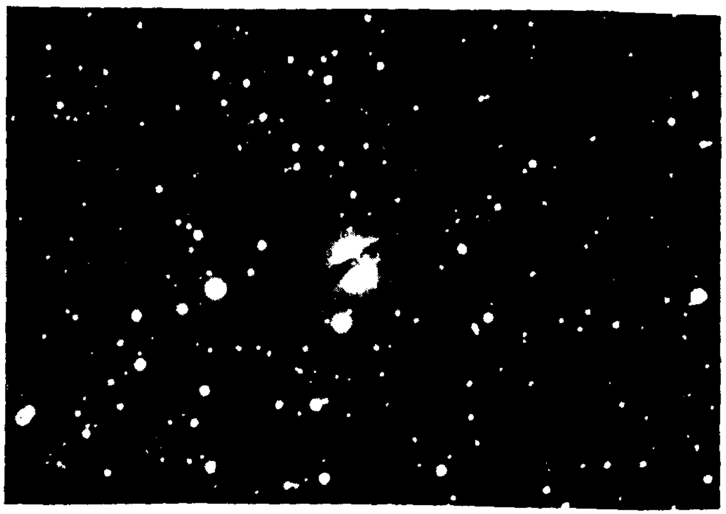


25. A modern radio-telescope in use near Sydney, Australia. It is fitted with a parabolic aerial reflector of dimensions 16 feet by 18 feet and can be used at wavelengths of 1½ metres, 50 centimetres or 25 centimetres.

26. A typical spiral nebula (Messier 81).



27. The Crab Nebula, which coincides in direction with that of a radio star in the constellation Taurus, the first to be identified.



28. Another likely radio star, the extra-galactic nebula NGC 5128. (*Franklin-Adams photograph.*)

FOREST ENTOMOLOGY

29. Laboratory of Insect Pathology, Sault Ste Marie, Ontario.

30. Canadian Forest Insect Survey. Rangers collecting specimens on ground sheet.

31. Defoliation by Spruce Bud-worm.

32. Method of collecting population samples from high trees.

further set of data which we shall call the number-magnitude data.

A fourth type of measurement has been performed on the nebulae and consists of measuring the relative amount of energy radiated in the blue and the yellow parts of the spectrum of a nebula. For this purpose, spheroidal nebulae are chosen because they are believed to form a more homogeneous group than do the other kinds. The relative energy is expressed by means of a number called the colour-index, C, and 26 nebulae have been investigated. The red-shifts for these nebulae have also been measured and this set of pairs of values of C and S we shall call the colour-displacement data.

DISTANCE

Even a superficial examination of the run of N against m in the number-magnitude data reveals that the number of nebulae increases in some kind of regular way with increasing m and that this regularity does not depend on the direction in which we look. Nevertheless, striking groupings of nebulae, called clusters, are observed in certain regions of the sky. A cluster may contain many hundreds of nebulae which evidently form a physically connected system. But the counts which provide the number-magnitude data run to millions of nebulae, whereas clusters contain numbers of the order of hundreds. Thus the clusters, though very important entities as we shall see, may be regarded as local irregularities in the nebular system that do not mask an overall regular increase of number with apparent magnitude.

In order to investigate this property of the nebular system in detail, it is convenient to replace apparent magnitudes by distances. In both classical mechanics and in general relativity, the fundamental kind of distance contemplated is metrical distance, which means the distance as it would be measured by ordinary measuring rods. Unfortunately, metrical distances in astronomy cannot be directly measured, even when the diameter of the earth's orbit is taken as the base-line, beyond about 300 light-years from the sun. In the nebular system all distances are luminosity-distances, D, by which is meant a distance computed by comparing the presumed intrinsic brightness of a nebula with its apparent brightness. Intrinsic brightness is measured by the apparent magnitude which the nebula would have if it were located at 3·26 light-years from the sun, and is denoted by the 'absolute magnitude' M. Thus to find the luminosity-distance D of a nebula two items of information are required; namely, an estimate of the absolute magnitude, which can only be found empirically, and secondly, a theory of the way in which the intensity of a light-source falls off with D. With regard to the second item, it is calculated that the intensity falls off inversely as the square of D and the calculation also shows that, for objects relatively speaking close to the sun, metrical and luminosity-distances are identical, but that, as fainter nebulae are considered, this simple equality breaks down. As to the absolute magnitudes, a method of finding them without appealing to distance-criteria was discovered in 1912. In the southern sky there are two nebulae called the Magellanic Clouds which are easily resolved into individual stars by telescopes of moderate power. They contain many stars, of the type called Cepheid after a specimen in the constellation Cepheus belonging to our own Galaxy, whose brightness rises and falls in a regular and easily recognizable manner. It was found that the Cepheids in the Small Magellanic Cloud exhibited a close correlation between the period of their light-oscillations and their mean apparent magnitudes. The periods varied from fifteen hours to one hundred days and the mean apparent magnitudes also varied over a wide range. Since the Small Magellanic Cloud is a faint object, it is reasonable to assume that it is so far from the sun that its overall dimensions are small compared with its distance. Hence differences of apparent magnitude amongst the Cepheids it contains may reasonably be attributed to differences of absolute magnitude, and not to differences of distance. It follows that the period of light-oscillation of a Cepheid star must be correlated with its mean absolute magnitude. Assuming that all Cepheids possess this property, the coefficients of the correlation can be calculated by finding the metrical distances of Cepheids which belong to our Galaxy. When this has been done, a 'period-luminosity' curve is

set up from which the distance of any newly-discovered Cepheid can be read off as soon as the period of its luminosity-oscillation has been determined. In this way it was found that the luminosity-distance of the Small Magellanic Cloud was 75,000 light-years and distances for other nebulae within one million light-years of the sun were similarly calculated from the Cepheid stars they contain.

For nebulae at greater distances, individual Cepheids can no longer be identified, but a study of the nearer nebulae shows that there is an upper limit to the absolute magnitude of the brightest stars in them and also to the overall brightness of an entire nebula. When therefore we observe a *cluster* of nebulae, we may assume that the brightest of the members of the group have, on the average, an absolute magnitude equal to the maximum absolute magnitude found for these nearby nebulae whose distances are known. The accuracy of such determinations is, however, probably not greater than 20 per cent. In this way it is estimated, for example, that the cluster of nebulae in the constellation Virgo is at $7\frac{1}{2}$ million light-years. The most distant clusters for which red-shifts have been measured have apparent magnitudes equal to about 18 and this is found to correspond with a luminosity-distance of some 250 million light-years, whilst the faintest nebulae included in the number-magnitude data are at 470 million light-years ($m = 21{\cdot}03$).

The formula for luminosity-distance in terms of apparent and absolute magnitudes is

$$\log D = 0{\cdot}2\,(m - 2S - M) + 1{\cdot}513 \qquad (1)$$

where $2S$ represents the correction to m arising from the red-shift, and D is measured in light-years. The displacement-magnitude data may easily be converted into displacement-distance data by using this formula to calculate the value of D for each m for which red-shifts have been measured. But this is not true for the entirety of the number-magnitude data because no values of S for m greater than about 18 have been measured. However, the number-magnitude data which refer to values of m less than or equal to 18 can be converted into a relationship between number, N and luminosity-distance, D, through formula

(1), since, in this range of m, the red-shifts, S, have been measured. When this is done it is found that N is approximately proportional to the cube of D which indicates that the distribution is approximately uniform in space. It is also found that the displacement-distance data can be interpreted as indicating that the red-shift is approximately proportional to D.

VELOCITY-DISTANCE RELATION

When the spectra of certain classes of stars of our own Galaxy are studied, displacements of the spectral lines are observed which are similar in character to the red-shift in the nebular spectra, except that, for the stars, shifts towards the violet as well as towards the red are noted. In such cases the shifts are attributed to relative velocities of the stars and the observer, a violet shift indicating a velocity of approach, a red shift a velocity of recession. This conclusion is based on the *theory* of the motion of light-waves and the shift is called a Doppler effect. For the case of red-shifts, the theory indicates that the velocity of recession corresponding to a shift of amount S is V, where $V = cS$ (c being the velocity of light). If we believe that the object of science is to group as many disparate phenomena as possible into one theoretical scheme, it is natural to interpret the nebular red-shifts as Doppler effects also. Hence the displacement-distance data may be expressed as a relation between the corresponding velocities V and the luminosity-distances D. In the first approximation, the linear relation between S and D is converted into a proportionality between V and D, the so-called velocity-distance relation. It leads to velocities which are very large, even by astronomical standards; for example, the red-shift for a nebula at about 250 million light-years corresponds to a velocity of recession of about 24,000 miles per second. But it must be remembered that these velocities are not directly *measured*; they are a product of the theory of the Doppler effect combined with the observed values of the red-shift. In principle we could refuse to interpret the red-shifts by means of the Doppler formula and introduce some new *ad hoc* interpretation of these spectral displacements.

EXPANDING UNIVERSES OF GENERAL RELATIVITY

If the velocity interpretation of the red-shifts is accepted, it is next required to interpret the empirical displacement-magnitude, number-magnitude and colour-displacement data in terms of a theory. To do this, an appeal to the laws of mechanics is required, and the particular system of mechanics we shall select for consideration is the theory of general relativity. We make this selection because general relativity has provided us with the most accurate description of gravitational phenomena which we possess and therefore we should like, if we can, to incorporate the observations made on the nebular system into this same scheme of thought.

Each nebula of the nebular system is an agglomeration of matter, whilst the nebulae are separated from one another by, as far as we can tell, empty space. By and large, there appears to be a uniform *distribution in space of the nebulae* around the Galaxy and these objects appear to be receding from us with speeds roughly proportional to their distances. Now in mechanics it is easier to deal with a continuous homogeneous distribution of matter than with a distribution consisting of discrete masses. Hence, 'model universes' are constructed in each of which the material is smoothed out to form a continuous homogeneous cloud of matter similar to a gas. Such a gas is characterized by a density ρ and a pressure p. It must be emphasized however that, in general relativity, such homogeneous model universes are to be regarded as first approximations only and not as necessary predictions of the theory. When a mathematical theory adequate to deal with discrete masses has been developed, it will be possible to construct models more in keeping with the observed small-scale irregular and disconnected arrangement of the nebulae. A second property that is assigned to the model universes is based on what may be called the 'modesty' postulate and is not, by its very nature, observationally verifiable. Astronomers are unwilling to believe that the Galaxy occupies a peculiar position in the nebular system, and the models are therefore constructed so that an observer situated at *any* point in one of them would observe the same kind of phenomena as we do from the Galaxy.

Working out the consequences of these assumptions within the general framework of Einstein's relativity, it can be shown, firstly, that there is a unique universal time t in each model such that ρ and p depend on t alone. Thus, at each instant, the density and pressure of the material are the same at every point of the model, and, in this sense, its spatial distribution is uniform. Secondly, it turns out that there are an infinite number of such models because the theory leaves three quantities undetermined, viz. – *firstly*, a quantity R which varies with t alone, secondly, a constant k, and thirdly the cosmical constant λ. If we knew the exact way in which R varied with t, we should have an item of great value: for R determines the density ρ and pressure p of the material as well as the velocity-factor h and the acceleration-factor f, these being quantities that indicate how the material is moving. But unfortunately the theory merely tells us that R depends on t – without, however, giving us the formula expressing this dependence. The constant k may be shown to have one or other of the values 1, 0 or —1 according as the volume of space at each instant t is finite ($k=1$) or infinite ($k=0$ or —1). As to the cosmical constant, it may be said that a close examination of general relativity shows that it is a mere constant of integration whose value must, however, be very small if gravitational phenomena such as are observed in the solar system are to be accounted for satisfactorily. The interpretations sometimes given to this constant by writers such as Eddington are additions to general relativity which may be questioned.

It is at this point that the views of theoretical astronomers diverge. One school of thought regards it as legitimate to determine the model through *a priori* assumptions which fix some or all of the three unknowns R, k, and λ. For example, it may be asserted that the volume of space is finite ($k=1$) at each instant t, and that the total mass of material in the model is also finite. Or a particular positive value may be assigned to λ which is related to the initial value of R, and so on. This method has been used by, amongst others, Einstein, Eddington, and Lemaître and it leads to a clear-cut choice of model. The other school of thought believes that the three unknowns must be fixed, as

accurately as may be, through an appeal to the empirical data and that *a priori* assumptions are illegitimate. Hubble, O. Heckmann, and the present author belong to this observational school which does not aspire to finality but only to discovering what model best fits the data. Indeed it is arguable that the definite results of the *a priori* school are due, not so much to the nature of the observed nebular system, as to the assumptions which have been superimposed on the data.

MODEL UNIVERSE AND OBSERVATIONAL DATA

It would be fair to say that both sets of investigators do agree about two things: firstly, it is impossible to construct a model which is co-extensive with the nebular system alone. The model always turns out to be larger than the nebular system, the additional parts being, because of the way the model is constructed, similar to those which represent the nebular system. Secondly, the value of the pressure p in the model must be zero *at the present moment* if the nebular system is to be correctly represented. Physically this pressure may be interpreted as a measure of the random motions of the nebulae over and above the velocity of recession. In the nebular system these random motions are observed to be small at the present moment, but this does not imply that they always were small in the past. Hence the model may have to take account of non-zero p as we recede in time.

The observational school proceeds to find the model which best fits the observations in the following ways: It is possible to define in the model quantities which correspond to the red-shift S, the number of nebulae N, and luminosity-distance D. Moreover formulae can be set up which express S and N in terms of D. Since, however, apparent magnitudes, and not luminosity-distances, are measured, it is necessary to link D with m, and this is done by adopting formula (1) as the connexion between these two quantities. Comparison of the theoretical formula connecting S and m with the displacement-magnitude data yields two items of information, viz. the numerical values of the velocity-factor h and of the acceleration-factor f for the nebular system. In the units employed, h and f turn out to be very small numbers indeed,

h being of the order of 10^{-10} and f of the order of h^2. As we have already remarked, h and f are connected with the unknown quantity R, and indeed they depend on the precise way in which R varies with t. Thus a knowledge of the numerical values of h and f does throw some light, though unfortunately very little, on the exact dependance of R on the time. The velocity-factor h is sometimes called the Hubble 'constant', though no proof has ever been given that it is not, as the theory indicates, a time-variable. At any rate, the numerical value of h corresponds to a velocity of recession of the nebulae which increases by 164 ± 3 km/sec for every distance-increment of a million light-years. The acceleration-factor f is very inaccurately determined and may indeed have the value zero.

Comparing the theoretical formula connecting N and m with the number-magnitude data and using the values of h and f already found, it is possible to conclude that $k=1$, and to obtain the numerical value of R at the present moment, as well as an estimate of the magnitude of λ. The conclusion that $k=1$ means that space is finite at each instant t. By this is meant that, though space is unbounded, yet it contains at each instant only a finite number of cubic light-years. Theory shows that this finite volume is $2\pi^2R^3$ which, incidentally, provides us with a physical interpretation of the quantity R as a number whose cube is proportional to the volume of space. The number-magnitude data show that R may have, at the present moment, any one of the values lying between 200 and 110 million light-years; but R is, of course, increasing with the time in a way given by h and f. It has sometimes been argued that R is too small, even if the extreme value of 200 million light-years is accepted, but this conclusion seems to be based on a misapprehension of the physical meaning of R. It is not a distance to a 'boundary' in the model, but merely a parameter which determines the volume of space at the present time. The part of this volume occupied by the nebular system can be calculated to be about 17 per cent of the whole, which is surely a small enough fraction of the total. The value of the cosmical constant λ is positive and is a small number of the order of h^2.

A further conclusion that can be drawn from the values of h, f, R, k, and λ, is that the average density of matter in the model is about 5×10^{-27} gm/cm^3. This density is rather high because estimates of the average masses of the nebulae suggest an average density of about 10^{-29} gm/cm^3 throughout space. The explanation of this discrepancy is yet to seek, but when we consider the many simplifications made in constructing the model, and the uncertainties of the data, it is surprising that any plausible results emerge at all. Unfortunately also, the model is very sensitive to small changes in the values of the data; for example, changing the red-shift correction in formula (1) from $2S$ to about S turns the model from one with a finite volume to one in which the volume of space is infinite. Greater certainty can come only through more and better data which may be confidently expected after the 200-inch telescope has been in use for a few years.

TIME OF EXPANSION

If the model universe is expanding in this way, it may be asked: How long has the expansion been going on? Has there not been an instant in the past at which the volume of the whole model was much smaller than is the volume now occupied by the nebular system? The observational school can give no definite answer to this question, because to answer it we must know exactly how the quantity R varies with the time, and this the data do not tell us. If, however, we take a leaf out of the investigations of the *a priori* school, and assume that $R=At$, where A is a constant, we can calculate from the value of h that the expansion has been proceeding, in exactly the same way as it is now going on, for some 2×10^9 years. At this date in the past, however, all the nebulae would have been concentrated into a zero volume and it is more than doubtful whether the assumptions made in constructing the model universe would still be valid.

The time taken by the light from moderately distant members of the nebular system to reach the Galaxy is an appreciable fraction of this age of 2,000 million years. Hence it might be expected that a comparison of distant nebulae with nearer ones of the same type would reveal secular changes. Now it is possible in the model universe to obtain a formula connecting colour-index

C with red-shift S and to determine the constants in this formula by using the colour-displacement data. These admittedly refer to 26 nebulae only, 8 of which lie at various distances between $7\frac{1}{2}$ and 240 million light-years. Introduction of the empirical data into the theoretical C—S formula does suggest that a secular effect exists. It turns out that the distant spheroidal nebulae have a rather different spectral energy-distribution compared with near-by nebulae. On the whole, the distant ones prove to be redder than they ought to be, even after allowance has been made for the effect of the red-shift. The reason for this change of colour is yet to seek but, for the moment, astronomers are concerned with establishing the effect observationally since measurements on 8 nebulae can hardly be regarded as providing a firm basis for any theory.

ALTERNATIVE THEORIES

The *a priori* method in cosmology is not restricted to the selection of one out of the many possible models in general relativity; it may be applied to the construction of the fundamental system of mechanics itself. We will refer briefly to two general theories which may be regarded as alternatives to general relativity. The first is called kinematical relativity and was propounded by the late Professor E. A. Milne, of Oxford. Kinematical relativity dispenses with the non-Euclidean geometries of general relativity and makes use of the Euclidean space and the mathematical structure of special relativity. The model universe it employs is by definition homogeneous and the fundamental kind of distance is neither metrical nor luminosity-distance, but one based on the notion of time. The distance of an object from the observer is proportional to the time-lapse, on his clock, between the moments of departure and return of a light-signal that has travelled to the object and back again. Moreover the observer's clock may work at any arbitrary rate because there must not be a unique uniform time in the model. A consequence of this definition of distance is that the object may or may not have a velocity relative to the observer according to the rate at which his clock happens to be working. Thus the red-shift, though it is a feature of the model

universe, cannot be unequivocally attributed to a velocity of recession and in this way is regarded as different in character from all other spectral displacements.

The second theory is associated with the names of Bondi, Gold, and F. Hoyle, and may be called the 'creation of matter' theory. It presupposes that the red-shift is strictly proportional to metrical distance and that the contents of each unit volume of the model universe are the same at all times. Since the unit volume is expanding with time, it is necessary to assume that matter is being created out of nothing at each instant and at every point in the model in order to keep the model in a steady state. The created matter appears in the form of hydrogen atoms which then, by some process whose details are not given, condense into nebulae.

In neither of these theories has the same detailed attention been given to the interpretation of the observational data as in general relativity, investigators having contented themselves with rough and ready methods. Against kinematical relativity, it may be argued that the definition of distance it employs is different from that used by surveyors or astronomers, whilst against the creation of matter theory, it may be said that there is no sound evidence in favour of a steady-state universe. Indeed, the colour-magnitude data suggest the exact opposite, namely, that the spheroidal nebulae in each volume of space were once of a different constitution from their present one. More fundamentally, however, both theories are, in the present writer's view, unscientific because, if they are accepted, it is necessary to suppose that the mechanics of the nebular system are in some way different from the mechanics of all other astronomical systems. If the object of science is to unify phenomena into theoretical systems with as wide an amplitude as possible, the general relativity interpretation may be accepted until it leads to some prediction seriously contrary to observation.

IDENTICAL TWINS IN DAIRY CATTLE IMPROVEMENT INVESTIGATIONS

JOHN HANCOCK

In the last half-century, the balance of animal production has been shifting from countries in which hand-feeding of animals is practised to countries where the animals themselves gather their feed. As a heritage of the past, it has been customary to consider the problem of improvement from two separate points of view: breeding and feeding. To-day, this attitude has largely changed, and it is generally recognized that the free-grazing animal and its environment cannot be considered as separate entities. When the environment has been improved to the economic limit, animals must be bred which best suit the available conditions.

There are reciprocal effects: those of the animals on their environment. They are seen in New Zealand in the deterioration of a large part of the hill country through over-grazing, and in the improvement of flat and undulating land through the fertilizing action of grazing stock. Both cases call for investigation; and, for efficient husbandry, for the use of the animals best fitted to the new conditions..

Research on cattle has been hampered by the lack of suitable experimental animals. This applies especially to dairy cattle, since the numbers in an experiment cannot be increased sufficiently to compensate for the inherent variability of the available material.*

* Certain devices have been adopted to obtain significant results from variable material, but these have rather severe limitations in interpretation, particularly in relation to pastured animals. Ecological surveys, though going a long way to clarify some of the principles which rule the inheritance-environment complex, can rarely guide development without support from experiment.

It might seem the obvious solution to eliminate variability by intensive inbreeding, as has been done with small animals for use in laboratory experiments. But it would take 10 generations, or at least 40 years, to obtain a line of cattle with an average inbreeding of approximately 87 per cent, and many lines would have to be started to obtain one which would survive. It would be an almost impossible task to develop several such lines, each representing a different level of inherited production-capacity, as would be needed for work on the inheritance-environment complex. This is fortunately unnecessary, since the equivalents of pure lines occur naturally in identical twins.

The occurrence of human identical twins has been known for a long time, and their potential usefulness in genetic research has been widely recognized. The principle of the method is as follows. There are two types of twins: 'identical' (*monozygotic*, uniovular or one-egg twins) and 'fraternal' (dizygotic, diovular or two-egg twins). Twins of the first type are the result of the splitting of a single fertilized ovum and are thus genetically identical. Those of the second type develop from separate ova, and are no more similar genetically than are full sibs (that is, than pairs of brothers and/or sisters). Any differences in physical or mental traits between the members of monozygotic twin-sets can be ascribed to environmental influences entirely, whereas differences between dizygotic twins are due partly to environment and partly to inheritance. A comparison of the magnitude of the differences found within the two classes of twins provides a direct, precise test of the relative influence of nature and nurture.

The occurrence of identical twins in cattle was disputed for a long time. Although a number of well-authenticated cases were sporadically described from 1922 onwards, it was not until 1932 that Knonacher, of Berlin, established their occurrence beyond doubt. However, in the following ten years the workers of the Berlin school dissipated their efforts on methods of recognition, and it is to Professor Bonnier, of the Animal Breeding Institute at Wiad, in Sweden, that the honour must go of having first demonstrated in practice the usefulness of identical-twin cattle as experimental material. From 1937 onwards, Bonnier was able to locate a sufficient number of such twins to start a series of experiments, and many important papers have since been published.

When on a visit to Sweden in 1938, Dr C. P. McMeekan became impressed with the possibilities arising from Bonnier's pioneer work. On his return to New Zealand, he initiated a search for identical twins, but owing to war-time difficulties and pressure of other work, he met with only little success and the attempt lapsed. However, when, in 1943, he became the Superintendent of Ruakura Animal Research Station, New Zealand, the project was taken up again and made the present writer's responsibility.

Some of the past indifference to the possibilities of twin-work was due, no doubt, to the feeling that identical cattle twins were so rare and so difficult to recognize as to be not worth locating. On the surface these misgivings seemed well founded. It has been calculated that only some 5 per cent of all cattle twins are identical, compared with 33 per cent of all human twins. As the overall twinning rates in cattle and man are equally low (about 1 per cent of all births) it would seem that the possibility of finding a sufficient number of identical cattle twins was rather remote. However, in a country such as New Zealand, where the dairy cattle and human populations are about equal there are, nevertheless, ten times as many identical dairy cattle twins born annually as there are human identical twins. This is due to the fact that almost the whole of the cattle population consists of females breeding regularly every year. In consequence, it has not been an unduly heavy task for the Ruakura Animal Research Station to locate nearly 300 sets of identical cattle twins during the last 6–7 years.

DIAGNOSIS OF IDENTICAL CATTLE TWINS

The diagnosis of identical cattle twins also proved easier than had been generally expected. As early as in 1932, Kronacher established the principles ruling the diagnosis of zygosity. He employed the 'similarity' method evolved by investigators of human twins. Recently, it has been shown by the present author that this method, slightly altered, can yield precise results. The

test applied at Ruakura consists of a critical although entirely subjective comparison of all visible characteristics, with the aim of providing an answer to the question: Are the differences of such kind, and is their sum sufficiently small, to justify the assumption that the twins are identical?

It is necessary to stress that differences are looked for rather than similarities, since all calves of a breed are superficially so similar. However, if a number of calves are examined in a critical frame of mind, it soon becomes apparent that they differ in a number of traits which make them stand out as individuals.

There are numerous characteristics in calves which lend themselves to easy visual comparison. Nose patterns, hair whorls, pigmentation of muzzle, mouth and ears, coat colour and the shape of head, lower jaw and muzzle have been found useful. Identical twins show generally only small or no differences in most of these characteristics, but may vary more in some traits than in others. For instance, identical twins may show easily recognizable differences in the extent of white hair spots (Inset, 7–10) but none in the shade of red or brown hair colour. In general, it can be said that identical twins show very small differences in characteristics which are bi-symmetrically expressed in an individual animal.

To check on the usefulness of various criteria the following test was carried out at Ruakura: 100 sets of fraternal twins were classified according to the number and kind of criteria on which they could be diagnosed as definitely non-identical. The difference-criteria used, and the percentages of cases diagnosed correctly on each, are shown below:

Hair colour	91%	Colour of eyelashes	39%
Muzzle shape	91%	Hair whorls	39%
Shape of head	72%	Pigmentation inside ears	36%
Shape of lower jaw	72%	Length of tail	25%
Body conformation	52%		

It will be seen that 91 out of 100 sets of fraternal twins showed sufficiently great differences in coat colour to make it possible to decide on the basis of this criterion alone that they were not identical. The shape of the muzzle was an equally good criterion, but the others also were very useful; 56 per cent of the sets could actually be diagnosed as fraternal on differences of either colour of eyelashes or pigmentation inside the ears or both. The great majority of fraternal twins differed in several criteria, 96 per cent in at least three, and 50 per cent in at least six. In these circumstances, it is easy to understand that a trained observer can recognize most fraternal twins in a few seconds. Only a very few mistakes have been detected when the original decisions have been checked against the red cell factor test evolved by American research workers.

UNIFORMITY TRIALS

Unlike previous workers in this field, it was decided at Ruakura to obtain a measure of the uniformity of identical twins reared under similar conditions before using them as experimental material. There were several reasons for this. One is that it is very difficult under wholly pastoral conditions, as in New Zealand, to eliminate all other environmental influences than those superimposed by the experimental treatment. In other words, not easily controllable environmental factors may cause appreciable errors. Other factors such as pre-natal environment, chance, and the process of egg-cleavage, could also play a part in lowering the efficiency of identical twins compared with animals selected at random. The influence of the sum of all these factors can be measured by the relative magnitude of the differences between members of the same sets of uniformly treated twins. As the numbers available at Ruakura made it possible to employ twins almost exclusively of other animals as experimental material, it was so much the more necessary to explore thoroughly their relative efficiency before committing ourselves, to such a drastic change-over.

Uniformity-trials have been carried out on the following characteristics: production of milk, butterfat, and casein; growth; grazing behaviour; physiological and biochemical blood components; physiological reactions to high temperatures; parasite burden and immunological responses. In addition to these, Mr J. James, of Ruakura, has used a number of identical bull-twins for uniformity trials on sexual strength, and Miss D. E.

Walker is at present making a very exhaustive study of the uniformity of anatomical development using an age series of identical twin heifers. Finally, Dr L. Wallace is studying the uniformity of feed-intake and digestive ability, using sixteen sets of free-grazing, lactating, identical twins. It is not intended to go into details of all these trials, but rather to give a general idea of the degree of concordance shown by identical twins under New Zealand conditions. Only the uniformity trials on butterfat-production, growth, and grazing behaviour will be described.

Butterfat production. Nine sets of identical twins were available for comparisons of butterfat production. At calving-time there were some unavoidable differences between individuals in age and date of calving, but these differences were no greater within sets than between them. This is mentioned because these factors may easily influence production-results, especially under New Zealand conditions where seasonal fluctuations in feed-supply may be an important source of error in such comparisons.

The figures in Table I show the production of butterfat taken to a standard 252 days' lactation.

TABLE I: *Butterfat Production in Uniformity Trials*

Twin No.	Butterfat lb.	Difference lb.	Twin No.	Butterfat lb.	Difference lb.
*T*1 *T*2	266 250	14	*T*3 *T*4	240 248	8
*T*11 *T*12	222 216	6	*T*13 *T*14	212 211	1
*T*17 *T*18	323 315	8	*T*23 *T*24	11 5	6
*T*27 *T*28	216 221	5	*T*29 *T*30	242 245	3
*T*31 *T*32	17 17	0			

Two facts are immediately evident: first, that the productions of the same sets were very similar; and, second, that there were great differences in productive ability between sets. The average within-set difference was only $5\frac{1}{2}$ lb. butterfat while the difference between the highest and lowest producing sets was no less than 311 lb. Even if the productions of the two very poor-producing sets were disregarded the range between sets would be as high as 107 lb.

The relative value of identical twins, compared with unrelated animals, can be expressed as the number of animals chosen at random which can be replaced in an experiment by one twin without loss of statistical efficiency. We call this the 'twin-efficiency value'. It is not a general figure, but refers only to one particular kind of experiment, the one from which it is derived. It depends on the following two factors: (*a*) differences between unrelated animals kept in the same conditions, and (*b*) differences between members of the same sets of twins. Both factors can be calculated from data of the kind presented for butterfat-production in Table I: (*a*) as the variations between sets of twins, and (*b*) as the variations within sets of twins. In these terms, the 'twin-efficiency value' is computed as the ratio:

$$\frac{\text{Variations between sets}}{\text{Variations within sets}}$$

For butterfat-production, we find the very high 'twin efficiency value' of 54. This is omitting the two very poor sets *T*23–*T*24 and *T*31–*T*32, the inclusion of which would have given a still higher value.

No matter how high the 'twin efficiency value', there are several reasons why more than one set of twins should be used in an experiment. One reason is that it is not possible to draw general conclusions from the reaction of one set of identical twins to a treatment, since they represent, after all, only one specific genotype. A second reason is that the 'error term', *i.e.* the part of the total variation that cannot be accounted for by differences in treatment, is likely to be substantially greater than is shown by uniformity-trials. This is because the experimental error cannot in

practice be separated from the treatment × set interactions which may be expected. Then again, there is the problem of wastage. If one member of a set of twins dies or becomes useless for experiment, it is obvious that in most cases her co-twin must be excluded. The wastage in twin-material tends, therefore, to be twice as high as when using unrelated cows.

For these reasons, at least five sets of identical twins are included in all experiment at Ruakura. In practice, it has been found that this number of twins is much more efficient in a statistical sense than even the greatest number of random animals that could be handled in an experiment.

Uniformity trials on growth. In a uniformity-study on growth, one sample comprised ten sets of twins which were reared together. At all stages of growth, the differences within sets of twins were small compared with the differences between some, at least, of the sets. This is illustrated in the weights of the 20 animals at one year of age shown in Table II.

TABLE II: *Growth in Uniformity Trials*

Twin No.	Weight lb.	Difference lb.	Twin No.	Weight lb.	Difference lb.
*T*1 *T*2	428 427	1	*T*17 *T*18	557 548	9
*T*3 *T*4	441 432	9	*T*19 *T*20	505 494	11
*T*7 *T*8	438 437	1	*T*23 *T*24	444 431	13
*T*11 *T*12	418 428	10	*T*25 *T*26	471 484	13
*T*13 *T*14	476 484	8	*T*29 *T*30	445 457	12

The average within-set difference was only $8\frac{1}{2}$ lb., while the difference between the heaviest and lightest animal was 139 lb.

An analysis of the growth-rates over seven periods, all beginning at four weeks of age, but varying in length from 12 to 84 weeks, provided an estimate of the usefulness of identical twins in experiments on growth. The calculated twin-efficiency value of 13 shows that identical twins are very valuable also in growth trials.

Uniformity trials on Grazing Behaviour. It has been rightly pointed out that the grassland agronomist always works with ecological communities – the sward complex and the animal complex – which in their two-way reaction make the accurate evaluation of pasture-yield a difficult task. The sward complex and the effects of the grazing animal on the sward have been investigated fairly thoroughly in the past, whereas the reaction of the sward (and other environmental conditions) on the grazing animal have received much less attention. While feed intake probably provides the best single index of a grazing animal's reaction to its environment, it seems certain that studies of the grazing behaviour of an animal can help to explain in rational terms its health and productive performance.

In order to gauge the usefulness of identical twins as experimental material in grazing-behaviour studies, a uniformity-trial was carried out involving six sets of lactating twins. Their behaviour was observed in eight 24-hour periods spread over the main part of the milking season. The observations were made in great detail, at ½-minute intervals, but for the sake of brevity only the average grazing times for each animal are reported here.

The figures in Table III show clearly that the time a cow spends in grazing depends to a large degree on her inherent individuality, as the members of the same sets of twins grazed within very small limits for the same lengths of time, whereas the differences between sets were in some instances very great. The average within-set difference was only 7½ minutes compared with the range of 137 minutes (or nearly 2½ hours) between the sets which grazed for the longest and the shortest times.

TABLE III: *Grazing Time in Uniformity Trials*

Twin No.	Grazing Time - Min.	Difference Min.	Twin No.	Grazing Time - Min.	Difference Min.
*T*1	403	4	*T*17	482	9
*T*2	399		*T*18	473	
*T*11	460	9	*T*23	338	5
*T*12	451		*T*24	343	
*T*13	387	13	*T*31	400	5
*T*14	400		*T*32	395	

The twin-efficiency value was very high, or over 70, indicating the marked superiority of identical twins over random animals for this type of work.

These few comparisons are sufficient to emphasize the superiority of identical twins over ordinary cattle. It is important to recognize, however, that twins are more useful in some types of investigations than in others. To take an example: identical twins are highly efficient in trials involving the measurements of the blood levels of magnesium or phosphorus but are hardly superior to randomly selected animals in regard to investigations on calcium or sugar levels. Information of this kind alone has fully justified the time and effort spent on uniformity-trials, as it is now possible to vary the number of twins used in an experiment according to the characteristic considered.

AN EXPERIMENT USING IDENTICAL TWINS

The easiest way of demonstrating the use of twins in experimental work on the nature-nurture complex is to describe an experiment which has recently been finished at Ruakura. The design of such an experiment can be explained most conveniently with the aid of the diagram in Fig. 1.

A total area of approximately twenty-six acres was divided into fourteen fields, or paddocks as they are called in New Zealand. Each paddock was divided into three breaks in the proportions 10 : 10 : 6. Ten cows were grazed in each break and moved into a new paddock every morning. Fifteen sets of identical twins were used and they were distributed as shown on the diagram. Five sets were split between *A* and *B*, five sets between *B* and *C*, and five between *A* and *C*. In this way, it was possible to obtain comparisons of equal statistical efficiency between the three groups of cows.

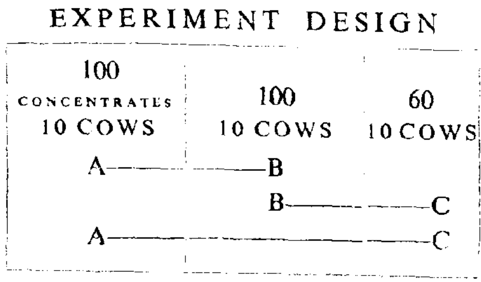


Fig. 1

In addition to pasture, hay and silage the cows in the *A* group received 1 lb of concentrates for each 5 lb of milk produced, irrespective of stage of lactation or level of production. They were thus fed in a way which should have made them produce to the limit of their genetical potentiality. The *B* cows were having the same amount of pasture on an area basis, but were not given concentrates. Hay and silage were fed *ad lib.* when pasture growth became inadequate, as happened regularly in winter and early spring and during the dry weather in summer. They were thus fed in a way which should have secured maximum production by the feeding of pasture and pasture-products alone. The cows in the third group, the *C's*, were fed similarly to the *B* cows but received only 60 per cent of the amount of hay and silage. Because they grazed an area which was only 60 per cent of that of the *B* group, it can be assumed that they were underfed, at least during periods of slow pasture growth.

The three treatments, *A*, *B* and *C*, have their approximate counterparts in practical farming conditions in New Zealand.

The *A* treatment represents the conditions on many farms where the main source of income depends on the sale of breeding stock. On such farms, it often pays to feed the cows to their genetical limit, even if this means that high-cost concentrates have to be used. The *B* treatment represents the good grassland farms on which the animals are kept on a high even level of nutrition at all times of the year, not by the use of concentrates but by the provision of ample pasture and an abundance of hay and silage when necessary. High production per cow is the aim under these conditions also, in the belief that this will lead to the greatest net returns. Finally, treatment *C* is representative of the great majority of farms, which generally are overstocked to a greater or less extent. The cows on such farms are consequently starved at certain times of the year, which presumably results in relatively low production per cow. The plan of this experiment was thus designed in such a way that the results should provide a direct answer to many pressing questions on breeding and feeding.

Results. The experiment ran for a period of three years. In order to avoid extensive tables, only the butterfat productions and the fat percentages obtained in the first year are presented and discussed. The results from the last two years were very much in line with those from the first year. The butterfat records of each of the thirty animals are shown in Table IV.

The great difference, 97 lb, between *A* and *B*, and 137 lb between *A* and *C* indicates that the New Zealand pastures and pasture-products are, from a quality point of view, not sufficiently good to sustain maximum productions. The difference between the two 'grass groups', *B* and *C*, was unexpectedly small. However, it must be remembered that all cows, even the *C*'s, calved down in similar, good condition. Furthermore, the amount of pasture per acre in the beginning of the experiment was similar for all groups. This meant that the *C* cows could, by grazing harder and for longer hours, obtain proportionately more than the 60 per cent intake they were allowed on an area basis. Observational grazing trials carried out for fourteen 24-hour periods throughout the season showed that the cows of

TABLE IV: *A – B – C – Experiment, Butterfat Productions*

A		B		C	
Twin No.	Butterfat lb.	Twin No.	Butterfat lb.	Twin No.	Butterfat lb.
T2	440	T1	312		
T17	512	T18	407		
T48	439	T47	319		
T49	237	T50	210		
T55	381	T56	278		
Average	402	Average	305		
		T14	212	T13	200
		T36	222	T35	193
		T37	380	T38	328
		T59	220	T60	222
		T63	294	T64	244
		Average	266	Average	237
T44	312			T43	168
T54	333			T53	219
T57	274			T58	203
T61	358			T62	107
T65	304			T66	197
Average	316			Average	179

the *C* group grazed on an average almost for one-and-a-half hours longer than their *B* mates and over two hours longer time than their *A* mates.

It seems, therefore, that the animals of the *C* group utilized the available pasture better than the cows of the other groups. Or to put it the other way round, the *A* and *B* cows wasted food, partly by being more selective grazers, thus leaving the more unpalatable grass untouched, and partly by using a portion of the ingested food for body-weight increments.

The average production, 359 lb for all ten cows in group *A* was surprisingly high, compared with the New Zealand average

of approximately 220 lb for stock of the same age. As the identical twins can be assumed to be a representative sample of the national dairy herd, this discrepancy in production is presumably due to differences in feeding levels. It looks, therefore, as if dairy cows in New Zealand are not fed in such a way that they could express fully their inherited capacity.

Just as the differences between the groups measure the effects of the environment, so differences within the groups provide an estimate of the effect of inheritance on production level. As can be seen from Table IV, the within-group differences are great. In the *A* group, there was a difference of 275 lb between the highest and lowest producer, in the *B* group 195 lb and in the *C* group 221 lb. Small age-differences between the animals tended to exaggerate these production-differences, but not to any very great extent. If the cows in any of the groups are ranked in order of production, it is found that the order of their co-twins in each of the other two groups is much the same. It seems, therefore, that inherited factors play a major part in causing variations in production within a group of cows. Peculiarly enough, this seems to apply in an equal degree both in groups of poorly and of well-fed cows. As a practical consequence, it seems that an animal which is a good producer in one environment will show a similar relative performance in other conditions, even if the environments are comparatively as extreme as the treatments *A*, *B* and *C* in this experiment. For this reason, it seems worth while to have inherently high-producing dairy stock, even if the feeding conditions are relatively poor. A further consequence is that it is almost as easy to select on production-performance under poor as under superior feeding conditions. Thus, this experiment does not substantiate the pedigree breeder's usual claim that he must feed his breeding stock at an extremely high and often uneconomical level in order to distinguish between good and poor producers.

Not all of the within-group differences were due to inherited factors; some must be ascribed to uncontrollable environmental factors or errors, and particularly to the interaction between environment and inheritance. A very striking instance of such interaction is provided by the performance of the twins *T*61-*T*62, which were split between *A* and *C*. *T*61, with a production of 358 lb of butterfat, was the highest of the ten animals in the *A*-*C* comparison, whereas her twin was easily the worst amongst the *C*'s, producing only 107 lb. The explanation of *T*62's poor performance is that she had an undershot lower jaw. In the beginning of the season, when the grass was fairly tall, she produced just as well as her sister, but as the grass became progressively shorter as the season went on, she was not able to collect enough food to keep up her production. *T*61 also had an undershot jaw, but because the *A* group was provided at all times with fairly long grass, and in addition was given concentrates, *T*61 was able to put up quite a creditable performance for a two-year-old cow (Inset 11). But apart from this exceptional instance of the interaction between environment and inheritance, there was not much evidence that this factor is important in causing differences between cows within a group.

For several reasons, it is important to discover to what extent the increase in the butterfat production of the experimental cows was due to an increased milk yield and to what extent to a higher percentage of fat in the milk contributed. The average fat percentages of each cow, together with the group averages, are shown in Table V.

A comparison of Tables IV and V shows that the differences within sets of twins were very much smaller in regard to the percentage of butterfat than in regard to the total quantity produced. It is thus clear that better feeding caused only a small increase in the butterfat percentages, and that the higher butterfat production of the better-fed cows was almost wholly due to their greater milk-yield.

This naturally raises the question of the relative importance of environment and inheritance. It is important to realize at the outset that there can be no clear-cut answer to this question if posed in general terms. The relative proportion of inherited and environmental influences on a given trait is not constant but varies in groups of animals according to (1) the genetic variability within the sample and (2) the range of environmental differences

TABLE V: *A* – *B* – *C* *Experiment, Butterfat Percentages*

A		*B*		*C*	
Twin No.	Butterfat %	Twin No.	Butterfat %	Twin No.	Butterfat %
*T*2	5.4	*T*1	5.3		
*T*17	5.7	*T*18	5.7		
*T*48	5.6	*T*47	5.3		
*T*49	4.8	*T*50	4.5		
*T*55	4.4	*T*56	4.2		
Average	5.2	Average	5.0		
		*T*14	5.2	*T*13	5.3
		*T*36	5.4	*T*35	5.0
		*T*37	5.1	*T*38	4.8
		*T*59	6.0	*T*60	6.0
		*T*63	4.5	*T*64	4.4
		Average	5.2	Average	5.0
*T*44	5.2			*T*43	4.9
*T*54	5.2			*T*53	5.0
*T*57	4.4			*T*58	4.5
*T*61	5.6			*T*62	4.9
*T*65	5.9			*T*66	5.6
Average	5.3			Average	5.0

superimposed on the animals. Thus inheritance will account for the greatest part of the total variation if the genetic variability is high but the environment is similar for all animals. On the other hand, environment becomes the greatest source of variation if the genetic variability is low, but the environment of each animal varies widely. An answer to the question, 'Which is more important, nature or nurture?' has therefore no meaning unless the population, and the environment in which each animal lives, are clearly defined. In the Ruakura twin-experiments, this condition is largely fulfilled. The twins, so far as is known, are a truly representative sample of the dairy-cattle population of New

Zealand at the time of collection, and the experimental treatments define in precise terms the environments superimposed on each animal or group of animals.

Now, the measured variations in given traits between the animals in a split-twin experiment, such as the $A - B - C$ trial described above, can be divided into three parts. The sources of variation are (1) differences between sets of twins, (2) differences between treatments, and (3) differences due to animal × treatment interaction and experimental error. The differences between twin sets express the variations due to hereditary effects, while the treatment differences naturally measure the effect of the environmental forces. The animal × treatment interaction term, which in an experiment of this type cannot be extricated from the experimental error, provides a measure of the degree in which the animals react, in their own specific way, to a given environment.

An analysis of the figures for butterfat-production and butterfat percentages in the $A - B - C$ experiment are shown in Table VI. The yield of the cow *T*62 with the undershot jaw has not been included, as inheritance × environment interactions of the magnitude shown by this cow are probably very rare.

TABLE VI: *A – B – C Experiment, Hereditary-Environmental Analysis of Butterfat Variations*

	Butterfat Production		Butterfat %	
	total variation	variation within treatments	total variation	variation within treatments
Hereditary portion of variation	47%	88%	90%	94%
Environmental portion of variation	47%		5%	
Heredity × environment interaction and error	6%	12%	5%	6%
Totals	100%	100%	100%	100%

It is obvious that butterfat-percentage is much more strongly inherited than is total butterfat-production. For this reason, it is clear that selection for manifested characteristics (phenotypic selection) is much more effective in changing the butterfat-percentage than butterfat-production. Precise knowledge of this kind may enable dairy cattle breeders to direct their breeding operations into the most effective channels.

The inheritance value obtained for butterfat-production is more than three times higher than previously reported figures. While, for several reasons, it is logical to expect that the twin method should give a high value, the 88 per cent obtained in this experiment was surprisingly high. The theoretical and practical implications arising from this finding are highly interesting but cannot be discussed here.

It is hoped that this article has indicated in a general way what can be done with identical cattle twins, and what has already been done. It is the author's firm conviction that identical twins open up new, exciting possibilities in almost any line of research on dairy cattle. They are especially useful to those who are interested in the interplay between environment and heredity. In fact, every properly designed twin experiment will yield valuable information on this problem, even though the primary aim of the experiment may be to measure the effects of various environmental factors.

FOREST ENTOMOLOGY

J. J. DE GRYSE

FOREST entomology, as a distinct branch of scientific forestry, had its beginnings in Germany during the early years of the nineteenth century. The successive trends of thought that developed in Europe were aptly summarized by S. A. Graham in 1939. He distinguished several more or less definite periods: the first, the natural history period; the second, a period of great taxonomic activity supplemented by experimental life-history studies; and the third, or modern, period in which the great emphasis is placed on experimental biology. He showed, too, that a parallel and somewhat comparable development took place in America, especially in the United States. Undoubtedly, the most important contribution made in the twentieth century consists in the shift of emphasis from taxonomy and old-fashioned natural history to the more precise disciplines of ecology and biometrics and their application to epidemiology.

The abundance or the scarcity of a species is regulated by two antagonistic forces: on the one hand, its ability to multiply; and, on the other, the power of the environment to limit such multiplication. The terms now generally in use for these two forces are 'Biotic Potential' and 'Environmental Resistance'.

The theory is axiomatic, but its application in experimental work is beset with numerous difficulties, not the least of which is the determination of the absolute egg-potential of a species. In practice, however, a modified concept of biotic potential, allowing for certain local or uncontrollable conditions, can be very useful, especially in appraising the influence of natural control factors on the progress of an outbreak, and in determining the nature and feasibility of artificial control measures. The factors entering into the calculation of the biotic potential are fecundity, number of generations, and sex ratio. The formula is $P(Rp)^n$,

where P is the original population, p the number of progeny per female, R the sex ratio and n the number of generations per year. Once the reproduction potential of a species has been determined, it is easy to calculate the mortality resulting from the destructive effects of environmental factors required to prevent the rise of a population. The following is a modified form of the formula devised by Bremer in 1928:

$$100q^n = \frac{100\,(a^n - b^n)}{a^n}$$

In this, q denotes the mortality quotient, a the average number of progeny per female, b the reciprocal of the sex ratio, and n the number of generations.

In general, such calculations show that an enormous proportion of the original population must be destroyed to maintain the species in a static condition, and that the required mortality quotient is directly proportional to the fecundity of a species and inversely proportional to the duration of its developmental cycle. The study of the action of individual or of combined environmental factors producing this mortality has, therefore, been the main concern of forest entomologists in the twentieth century. Whilst there is fairly general agreement regarding the nature of these factors, there always has been and still is a marked division of opinion on their relative importance in the economy of nature. During the first quarter of the present century, the belief prevailed, in many quarters, that the rise and fall of insect outbreaks were governed almost entirely by the degree of efficiency with which insect parasites and predators performed their regulatory function in the biotic community. These ideas are rapidly losing ground and are being replaced by a steadily growing realization of the importance of factors other than biotic. Many of the older workers were well aware of the fact that weather and climate played a weighty part in the origin and development of infestations. Their notions, however, were vague, being mostly based on casual observations sadly lacking in precision and continuity. The principal credit due to their modern successors lies not in the promotion of a new idea, but rather in the adoption of new methods. This viewpoint is aptly expressed by Escherich in *Die Forstinsekten Mitteleuropas:*

> What is new in our time, and what gives applied entomology a new front, is the attempt to endow with precision and to interpret, wherever possible, on a mathematical basis, those relations between climate and outbreaks which, heretofore, were only suspected, instinctively surmised or inadequately substantiated. This was accomplished by improved experimental methods, by the most painstaking, long-term and thoroughly planned field observations, and by historical-statistical researches covering extensive periods.

Under the leadership of such men as Janisch, Friedrichs and Bodenheimer, the study of bioclimatology received an unprecedented impetus. A few of the advances in concepts and methods to be credited to the new school, and now generally accepted in experimental biology, should perhaps be briefly mentioned: among the most outstanding are Friedrich's distinction between macro-, meso-, and microclimates; Pierce and Bodenheimer's zones of viability based on temperature-humidity relations; Blunck's theory of thermal constants and Cook's graphic method for the comparison of climates from region to region.

Some of the protagonists of climatic causation of insect outbreaks go so far as to regard all other factors as of minor or no consequence. On the other hand, some entomologists are inclined to lay special stress on nutritional factors. They argue that 'when there is an unlimited and convenient supply of a certain species of tree, then the stage is set for an outbreak of the insect pests of that tree' (Graham, 1939). Strictly speaking, such a state of affairs is a condition favourable to development of large-scale destructive infestations, but it cannot be considered as a true cause.

Before closing this chapter on general principles and methods, it seems fit to draw attention to the fact that it is characteristic of nearly all modern investigations to place great weight on a thorough-going mathematical analysis of data. This, in itself, is a tremendous advance over older methods in which quantitative measurements hardly ever progressed beyond the stage of arithmetical calculations of averages and percentages, without critical appraisal of their significance. Whilst the new method

does not place biological research in the category of the 'exact sciences' it inspires a degree of confidence in its conclusions that was never before justifiable.

Evidently, the experimental procedures, the theories, the principles and methods broadly outlined above are not the exclusive appanage of forest entomology, they are simply recent developments in the field of general entomology. The impact of these innovations on the practice of forest entomology will be discussed in somewhat greater detail in the subsequent paragraphs under the headings of surveys, epidemiology, and prevention and control.

FOREST INSECT SURVEYS

F. S. Bodenheimer in 1938 expressed clearly and succinctly the most important methods and aims of modern Forest Insect Surveys:

> All studies of animal communities must begin with quantitative counts of the animals present in small areas. The analysis of a sufficiently large series of such counts made in a certain region allows a comparison to be made between the species and their numerical abundance in each area and permits the description of the statistics of animal communities at the time of the survey. The study of environmental factors acting in each area, as well as the study of the physical ecology of any species, enables us to understand the dynamics of animal communities.

Surveys are to forest entomology what diagnosis is to medicine. Rightly or wrongly, it has been an accepted dictum that insect outbreaks, like forest fires, can be most effectively dealt with in their early stages and timely detection of such outbreaks has been the goal of forest entomologists the world over. Periodic inspection of woodlands by individual foresters, more or less quantitative appraisals of the abundance of individual insects of special economic importance, intensive ecological investigations of the insect populations on restricted areas have been undertaken from time to time in the past. It is only in recent years, however, that the scope of such explorations has been broadened so as to constitute, as far as possible, an annual census of the entire insect population of large forest regions. As such, they have become gigantic studies in synecology and are gradually leading to a much better understanding of the causes of insect epidemics.

Adapted, as they must be, to the needs of the region and the availability of resources and personnel, the systems in vogue in different countries vary greatly in organization and in methods. Perhaps one of the oldest and most elaborate surveys of this kind – not necessarily the best – is the Forest Insect Survey of Canada.

It is common knowledge that Canada's national economy is, in a large measure, dependent on the conservation of its forest resources. The ever-increasing demand of our civilization for wood products of many kinds constitutes a severe drain on the available supply. The frontier of the Canadian forest is receding rapidly before the axe of the exploiter. Fire, insects and diseases take a heavy toll of the resources left by industry. At the present rate of depletion, the Canadian forest can remain commercially productive for only a very few decades. Regulation and control of every agency of destruction, whether man, beast, or the elements, has become imperative. The enormous losses caused by insects were not fully recognized as a serious menace until now. Indeed, Canada, because of its geographic position in the northern part of the temperate zone, is more vulnerable to large-scale insect attacks than other parts of the American continent. Its climate puts a severe limitation on the number of tree species that can maintain themselves in the boreal forest. The result is seen in the enormous tracts of forest stands of uniform composition and often of even age. Whilst such a condition may not, in itself, be the cause of insect outbreaks, there can be no doubt as to its propitiousness for the wholesale destruction of thousands upon thousands of square miles of valuable timber so frequently recorded in the past. In the face of this constant threat, a thorough knowledge of present conditions and, if possible, a knowledge of 'the shape of things to come' are valuable assets in making plans for active war upon forest pests. From the economic standpoint, diagnosis and prognostication are the principal aims of the Forest Insect Survey. From the scientific angle, increased knowledge of certain phases of taxonomy, distribution, biology and cycles of abundance is the goal.

Complete attainment of these objectives, over a territory as vast as the Canadian forest, is obviously impossible. Nevertheless, the results achieved during the last thirteen years have yielded such an enormous amount of valuable information as to justify fully the effort and the money invested in the enterprise.

The Canadian Forest Insect Survey has its national centre in Ottawa, Ontario, and is headed by a Coordinator whose function consists in directing the activities of ten regional centres located at strategic points in the provinces from Nova Scotia to British Columbia. Each regional laboratory has a staff of scientists and technicians who are charged with laboratory and insectary work connected with the study of samples received from the field. Seventy-five trained forest insect rangers employed by the Government of Canada act as leaders in the collection of data and insect material. They are assisted in this work by some 3,000 co-operators belonging to industrial companies and provincial forest services. These men are stationed at ranger-posts scattered throughout the entire accessible forest and make regular observations, collections and reports on conditions in their respective territories. In 1936, the first year of operation, about 500 such reports and collections were received from the field men. By 1949 and 1950, this number had risen to over 30,000. The interest of co-operators is maintained by acknowledging immediately every collection received and giving the sender a tentative identification of the specimens contained in his shipment.

The greater part of the material sent in by the field personnel arrives at the regional laboratory in the larval or pupal stages. All, or nearly all of it is carefully recorded, tentatively identified and finally placed in rearing cages with a view to obtaining final determination of the adults, and information on the natural control agents, parasites and disease, which may be operating in the population from which the sample was derived. Overwintering larvae and pupae are placed for a time in cold storage at suitable temperatures and are 'forced' through in incubators during the later part of the winter. In this way, all the material collected in one season is disposed of before the beginning of the next.

Defoliating insects make up the greatest part of the field collections. They are mostly obtained by beating suitable trees with an axe or a pole and gathering all the insects that fall on a standard size mat spread beneath the tree. The specimens are placed in a box with a small amount of food and shipped to the nearest survey centre. Insects other than defoliators are gathered by different methods. With each shipment, the collector submits a report on a form specially prepared to show all the most important data relating to the insects and their habitat.

During the early years of operations the sampling was purely qualitative, and even to-day the contributions of the outside co-operators must still be regarded as such. At the same time, the work of the trained forest insect rangers is becoming progressively quantitative through the establishment of a network of thousands of permanent sample plots which are re-examined at regular intervals. Except in a very few special cases, the system is not sensitive enough to permit the detection of incipient outbreaks in the very early or 'preparatory' stages. Usually, however, the symptoms of the 'prodromal', or premonitory, stage can be recognized well before the 'eruptive' stage is reached.

The highlights of the Survey are published in an annual report which presents in brief popular form all the information likely to be of interest to the forest industry and the forest services. In addition, there is, in the files, a tremendous backlog of scientific data awaiting further elaboration and eventual publication in a series of technical bulletins.

Recording and analysing the hundreds of thousands of data accumulated over the years is an almost superhuman task. To facilitate these operations, a system is being devised for coding all information and entering it on punch-cards with a view to making use of business machines in its interpretation.

EPIDEMIOLOGY

The definition of forest entomology as 'that branch of the biological sciences which deals with insects in their relation to

forests and forest products', given by Graham in 1939, immediately conveys the idea that ecology must be the *leitmotiv* in most of the research activities underlying its practices. Etymologically, ecology is the 'science of the house'. Scientifically, it is the 'science of the environment'; the house is still there, but it has become metaphorical and is now inhabited by all sorts of creatures, either living together on friendly terms or fighting with each other for a place in the sun and sometimes for a place in the shade. Such a house is the forest, with its trees, its myriads of tiny organisms, its large and small vertebrates, its hot spots and cool spots, its wet places and dry places, its bright nooks and dark corners. The study of the interaction of all these living creatures and all these inanimate elemental forces is the fascinating task of the ecologist. Whether or not the cohabitation of all the living things in a forest community is more or less fortuitous or whether it has resulted in the formation of a unified superorganism is a question of considerable academic interest, but of no particularly practical import. The fact remains that there is considerable interdependence between many members and groups of the association and that the life-processes of all are profoundly affected by the physical factors of their surroundings. The net result of these harmonious or conflicting forces is expressed in more or less marked fluctuations in the numbers of constituent populations. The measurement of these fluctuations is the province of the epidemiologist, in this case, the forest entomologist. Taking the census of an insect population is by no means a simple assignment, especially in the natural forest. The vast areas involved, the enormous size of populations, the comparative inaccessibility and the extreme variability of the sampling universe, all combine to render the task complicated and subject to experimental error. Many methods have been tried with varying success. Direct counts of all the insects present on entire trees, on sections of the crowns, or on a number of measured twig samples have been attempted by several investigators. Others have used the degree of defoliation as an index of population density, still others have advocated the measurement of [illegible] drop (larval excrement), light traps and traps baited with sexual or chemical attractants have been tested. All these devices have their good points and their bad, and a preference for one or other must be determined by the objective. In general, however, especially in connexion with larval or pupal counts and estimates of defoliation, it is safe to say that large numbers of small samples give a better measure of populations than do small numbers of large samples.

Inquiry into the causes of the fluctuations presents an even greater challenge to ingenuity. This is particularly true in the case of primary insects, that is to say, insects that attack perfectly healthy trees and are able to complete their entire life-cycle on them. Secondary insects can live only on trees or parts of trees that have been previously weakened or killed and the agencies causing deterioration of the host are usually responsible for the ensuing gradation of the pests. The school of thought led by L. O. Howard and W. E. Fiske, of the U.S. Bureau of Entomology, held the view that the biological factors of the environment, more specifically insect parasites and predators, were the principal, if not the only regulators of the density of insect populations, whilst Bodenheimer and others attributed all fluctuations in insect numbers to the influence of climatic factors. In reality, the controversy was not quite so simple as that except, perhaps, in its initial stages. In later years, however, the contending parties have modified their views and a fairly reasonable compromise was reached, as shown by the following excerpt from Bodenheimer's *Problems of Animal Ecology:*

> The climatic school was right in pointing out that under normal conditions climate really controls population density in being responsible for the highest degree of destruction. It is, as a rule, the main factor as long as the saturation point of environmental capacity has not yet been approached. This is the normal condition for most animals and even for the longer span of life of species which are inclined to outbreaks. We were wrong to underestimate, in the infancy of our discoveries, the actual part played by biotic factors, especially during the breaking down of gradations.
>
> The biological school was right in making a distinction between density-dependent and density-independent factors. Only the former may be relied upon to bring about the crisis, provided a species in gradation, or in any other way, has approached or surpassed the local specific environmental capacity. Climate may do so, but cannot be relied upon. The school was wrong in limiting biotic factors to the parasite component solely. Parasites are only one of the biotic factors: predators, diseases, effects of population pressure, and lack of food are others. These biotic factors, acting as a whole, will end any outbreak, whereas the single factors are interchangeable. The special stress laid on parasites was unjustified.

Last year, a most interesting and potentially far-reaching contribution to bioclimatology was made by W. G. Wellington. In his studies of the epidemiology of the spruce budworm (*Choristoneura fumiferana* Clem.), this author found that there exists a good correlation between the annual number of cyclonic centres passing through restricted areas and outbreaks of the spruce budworm which are always preceded by a decrease in the numbers of these centres. Characteristically, precipitation falls below normal during pre-outbreak periods. Droughts in June and July are most common, but sometimes spring or autumn droughts are also associated with these summer dry periods. The inference is that widespread outbreaks of budworm are the reactions of permanent populations to a relaxation of climatic control in suitable loci, rather than the result of dispersal over long distances.

As previously mentioned, a third school of forest entomologists claim that nutritional factors, namely the superabundance and suitability of the food supply, are mainly responsible for the rise of infestations. This view is held by many workers in America, especially those who are engaged in studies of the spruce budworm problem.

A discussion of the causes of outbreaks would not be complete without a passing mention, at least, of the theory that outbreaks may be the effect of the mass survival of weak physiological strains that usually perish under normal environmental pressure, but are periodically favoured by a combination of aberrant physical or biotic factors (Franz, 1944).

Notwithstanding the vast amount of research on insect epidemiology undertaken within the last fifty years, the final answer to the question, 'What is the cause of insect outbreaks?' has not been forthcoming. It is more than likely that the solution of this problem will be found in a combination of causal relations and that this will be different for every species of insect. Possibly a more intensive study of insects under (normal) endemic conditions will be just as fruitful.

PREVENTION AND CONTROL

It is customary to classify the control measures used against forest insects in four principal categories: silvicultural, biological, chemical and mechanical. It seems more appropriate to consider silvicultural practices as preventive rather than control measures. Their objective is the creation of highly productive forests that will eventually be more or less resistant to insect attack and the prevention of the wholesale destruction of large forest stands. Regulation of the insect's food supply is the basic procedure. This is accomplished mainly by the adoption of a system of harvesting and reforestation aimed at producing suitable mixtures of tree species and differentiation in age-classes. M. L. Prebble has lately summarized the most important principles:

> General recommendations for preventing insect damage through cultural practices or through forest management are based on the concepts of maintaining stand vigour; controlling the abundance of component species and the distribution of age-classes; preventing the accumulation of overmature and senile stands; avoiding the planting of species unsuited to the soil and climate; and adhering to stand mixtures found in the climax forests characteristic of the site and region. *These recommendations, and variants of them, are predicated on the belief that in most cases good silvicultural practice will not only improve growth and yield, but will also tend to reduce the danger of destructive insect outbreaks.*

Ideally, at least, well-planned salvage and presalvage cuttings should be classified under cultural control measures. They should be aimed not merely at reducing immediate losses but at preventing, as far as possible, the development of future infestations as well.

There cannot be the least doubt that, of all the methods of combating forest insects, silvicultural management holds the greatest promise of permanency. In Western Europe, where most of the forests have been planted by man, this sort of protective silviculture can be most easily practised. In countries like Canada, with its natural forests, as long as sizable reserves of virgin stands are available for exploitation, operators will be reluctant to adopt any intensive forestry methods that may increase the costs of production. As the forests recede farther from the mills, this attitude is bound to change.

The reduction of insect populations by parasitic or predaceous organisms is generally known as biological control. Mammals, birds, insects, worms, protozoa, bacteria, fungi and viruses are the regulatory agencies. Enormous numbers of insects, passing a part of their life in the forest floor, are destroyed by small mammals. Birds attack them in all their developmental stages and the same is true of insect parasites and predators and of the various types of micro-organisms mentioned above. The relative importance of any of these groups varies with the insect host species. Whether biological control is as effective in regulating host population densities as it was thought to be some years ago, is a moot question. Although in the opinion of many, factors other than biotic contribute more heavily to the mortality quotient required to preserve the *status quo* of a population, it should not be forgotten that the comparatively smaller numbers of insects killed by parasites and predators may be just enough to bridge the gap between a state of equilibrium and that of a rising epidemic.

Until recently, the term biological control was applied almost exclusively to the use of insect parasites and predators for the control of insect pests. The method has been practised for many years, especially in the struggle against noxious insects accidentally introduced from foreign countries. American entomologists were the pioneers in this field. The first planned introduction of a predaceous insect into the United States was that of the ladybird *Novius cardinalis* from Australia to California in 1888 to combat the cottony cushion scale *Icerya purchasi*. The importation of parasites for the purpose of controlling a forest pest was first undertaken between 1905 and 1913, when more than thirty species of parasites were brought from Europe and Japan to fight the gypsy and browntail moths in New England. At the time of the great larch sawfly outbreak that swept over Eastern Canada from the Atlantic to the region of the Great Lakes, several colonies of *Mesoleius aulicus*, collected in England, were liberated in the provinces of Quebec, Ontario and Manitoba. Although this parasite became established, it made little or no impression on the sawfly outbreak of that day. However, in subsequent years, it repeatedly showed its value as a control factor by effectively repressing severe infestations of the sawfly in the new larch stands now replacing those that were destroyed during the first attack. As the present larch forests approach maturity it is probable that they will lose some of their resistance to repeated defoliation and that the protective value of the parasite will become materially reduced. To counteract such an eventuality, prompt utilization of larch as soon as it reaches merchantable size is recommended. This combination of cultural and biological methods should be the means of preventing serious losses in the future.

The introduction of *Mesoleius* marks the start of biological control of forest insects in Canada. Since that time, insect parasites and predators have been used in combating a number of other forest pests. The largest single operation of this kind was undertaken in the fight against the European spruce sawfly between 1931 and 1938. Although several parasitic species introduced at that time have become successfully established and are becoming increasingly effective, they played only a secondary role in bringing about the decline of the sawfly outbreak. A virus disease, which was first observed about 1936, and which spread with phenomenal rapidity over the entire infested territory, caused such tremendous mortality in the sawfly population as to reduce it to endemic levels in the short span of about four years. The disease is highly specific and, so far, all attempts at inoculating native species have resulted in failure. It seems probable, therefore, that it is of European origin and that it may have been accidentally introduced with shipments of the parasites. The virus is transmitted by mouth. Advantage has been taken of this fact and suspensions of the polyhedral bodies, sprayed on the

food plant, have been used in the artificial dissemination of the disease in Newfoundland, and in other new centres of infestation.

It is only natural that the spectacular decline of the European spruce sawfly outbreak should have served as a strong incentive to renewed interest in the possibilities of diseases as control factors in insect epidemics. A further stimulus was provided through recent, somewhat similar, occurrences in connexion with outbreaks of the hemlock looper and the black-headed budworm in British Columbia. Furthermore, recent advances in knowledge of viruses, bacteria and fungi have removed many of the objections and broken down many of the prejudices formerly held against the use of pathogenic organisms. The time seemed ripe for action, and research on the subject was undertaken in earnest. After a period of preliminary investigations, it was decided to embark on a formal programme of studies and to erect a special laboratory for the purpose at Sault Ste. Marie, Ontario.

This laboratory was completed in May, 1950. In its construction every possible precaution has been taken to prevent the contamination of the working space as well as cross-contamination of cultures. The glass walls, the terrazzo floors and the cement-enamel ceilings can be readily treated with disinfectants. Most of the work tables are built-in; they are made of cement covered with structural glass. All the air entering the building is sterilized by passing first through a precipitron, then under a bank of ultra-violet light and finally through a set of carbon filters. The air in the rooms and that in the rearing units is introduced and exhausted through separate ducts. All lights are controlled from a central switchboard. There are air-locks equipped with ultra-violet lights between the corridors and all the rearing rooms. The most modern equipment required for advanced research has been provided. The work is organized in three sections; virus diseases, fungous diseases and bacterial diseases. It is hoped that a section for the study of protozoan diseases will be added before long.

This laboratory will serve as a national centre for fundamental research and for the mass-propagation of pathogenic organisms.

Outposts for the study of epidemiology under field conditions will eventually be located at strategic points throughout the country. At the present time, special attention is being directed to studies of diseases of the spruce budworm, the European pine sawfly (*Neodiprion sertifer*), the European spruce sawfly (*Gilpinia hercyniae*), the forest tent caterpillar (*Malacosoma disstria*), and the gypsy moth. Important contributions to the knowledge of the multiplication of virus diseases have recently been published by Bergold (1950) and rather spectacular results have been obtained by Bird (1950) in the artificial propagation and dissemination of viruses attacking the European spruce sawfly and the European pine sawfly.

Mechanical and chemical control methods are the last resort of the forest entomologist faced with an emergency. They belong to the 'do or die' class. Mechanical control consists of such practices as banding trees, felling or girdling of trap trees, felling and burning of brood-trees, hand-picking of insects, etc. The successes that have sometimes been claimed for some of these methods are, to say the least, questionable and often greatly exaggerated. Chemical control consists mainly in the application of insecticides by means of ordinary power-sprayers or by means of various types of distributing devices mounted on aeroplanes. The discovery of the insecticidal properties of DDT and some other new compounds has given a great impetus to chemical control. It cannot be denied that the immediate results of a spraying operation with DDT are astounding. At the same time the indiscriminate use of such substances is fraught with real danger. In any event, the process is costly and its general usefulness is limited by topographical difficulties and meteorological conditions. In rare instances, however, as in the protection of extremely valuable trees or of infested stands that are to be cut in the near future, the application of insecticides may be a very effective way of preventing serious financial losses.

To sum up, it is safe to assume that, generally speaking, silvicultural and biological control methods offer the best solution of the majority of our problems. At the same time, it should be

realized that we can easily be led astray by our enthusiasms and that it is a serious but common mistake to 'put all one's eggs in one basket'.

SELECTED BIBLIOGRAPHY

Balch, R. E. 'Cultural Practices and Forest Insects.' *In* 65th Ann. Rep. *of the* Entomol. Soc. of Ontario for 1934. 'The Relation between Protection and Management.' *In* Proc. Soc. Am. Foresters Meeting at Boston, Washington, D.C., 1948.

Bergold, G. E. 'Multiplication of Insect Viruses as Organisms.' *Can. J. Res.* 'E,' 28, 5 (1950).

Bird, F. T. 'The Dissemination and Propagation of a Virus Disease affecting the European Pine Sawfly, *Neodiprion sertifer*.' *In* Bi-Monthly Progress Rept., Forest Insect Investigations, Canada, Dept. of Agr., Vol. 6, No. 5 (1950).

Bodenheimer, F. S. *Problems of Animal Ecology*. Oxford, University Press, 1938.

Escherich, K. *Die Forstinekten Mitteleuropas*, Pt. III. Berlin, Paul Parey, 1931.

Franz, J. 'Über die genetischen Grundlagen des Zusammenbruchs einer Massenvermehrung aus inneren Ursachen.' *Z. Ang. Ent.*, Bd. 31, Heft 2 (1944).

Graham, S. A. *Principles of Forest Entomology*. New York, McGraw-Hill, 1939.

Howard, L. O. and Fiske, W. E. Bull. 91, U.S. Bureau of Entomology, Washington, D.C., 1911.

Prebble, M. L. 'Forest Entomology in Relation to Silviculture in Canada.' *Forestry Chron.* (March, 1951).

Swaine, J. M. 'The Relation of Insect Activities to Forest Development as Exemplified in the Forests of Eastern North America.' *Sc. Agr.* **14**, 8 (1933).

Wellington, W. C. *et al.* 'Physical and Biological Indicators of the Development of Outbreaks of the Spruce Budworm.' *Can. J. Res.* 'D,' 28, 348 (1950).

PLANT NUTRITION AND DISEASE

PARASITE and host exist in a complicated relationship. For the parasite to survive, it must flourish, but not too well: it must not kill its hosts, or not too quickly. Alternatively, it must be capable of an independent existence in a host which is less vulnerable. Whether parasite or host will be favoured by changed conditions must depend on a fine balance between their respective requirements. The following note relates to a group of fungus diseases of plants – those producing Verticillium wilt in cotton, tomatoes, hops, and other species – in which improved nutrition of the host favours the parasite on balance. In other examples, the effect is reversed, the improved nutrition of the host enabling damage by the parasite to be more readily made good. In others, again, although the individual plant may benefit, so also does the parasite, with the result that infection is extended, though less severe.

Study of plant nutrition in relation to *Verticillium* fungi goes back to 1939, when the Rothamsted Experimental Station, Harpenden, was asked by the Empire Cotton Growing Corporation to undertake a general inquiry into the wilt disease which they cause. Infection is from the soil through the roots, and thence through the vascular tract until the whole plant is infected. The investigation was on tomato plants, grown under glass. Wilt was greatest in plants grown in soil with a high content of available nitrogen, and was also increased by a high intensity of light and by other conditions making for a quick assimilation of carbon dioxide by the plant. Related experiments have been carried out both with other fungi, and with the same fungus growing in other plants.

Verticillium wilt of hops, caused by the same fungus, *Verticillium albo-atrum*, which affects tomatoes, has been a serious

disease in Great Britain since 1932. It exists in a mild and a severe form, of which the latter is thought to have arisen as a mutant at a single Kentish farm about 1930, but has so far spread that more than 100 farms are now infected. From an initial yellowing of the lower leaves, the disease progresses until eventually the whole plant is dead. Control, at present, is through resistant varieties, though there is none yet which is wholly resistant. By 1949, sets of the first of these varieties had been distributed to 15 growers. The proportion of infected plants, both in this and another resistant strain, averages 2 to 3 per cent, and plants infected in 1948 seem since to have recovered – which increases interest in the factors which favour plant or parasite.

Preliminary nutritional experiments were carried out in 1944-7 in conjunction with the Long Ashton Research Station of Bristol University. These were with hop plants grown in sand cultures, supplied with various nutrient solutions. As with tomatoes, nitrogen-deficient plants showed much less severe symptoms than those grown with a complete nutrient solution, and the same conclusion has been confirmed by the East Malling workers for hops grown in nitrogen-deficient soil, compared with plants grown in the same soil but with extra nitrogen provided. Other experiments have been carried out on tomatoes, as at Rothamsted, and on antirrhinums and asters, which it is thought will be more satisfactory in their showing of symptoms. Again, the severity of infection is increased with a greater nitrogen supply to the plant.

Finally, two new methods have been devised by which quicker and more standardized infections can be produced. They have been described by Dr W. G. Keyworth, of the East Malling Research Station. With the smaller plants, which root more quickly, this has been done by standing cuttings in a suspension of fungus spores before potting; and locally in hops by injecting the fungus in a solution of agar jelly into the hollow stem internodes. Both methods are expected to be of wider application, with that used for hops suitable also for other hollow-stemmed plants. The first is being used already in the 'screening' of varieties for disease resistance, and in tests of chemical methods of treatment.

The nutritional problem has been discussed in more general terms by Dr F. C. Bawden, of the Rothamsted Station. The nearest to a generalization that is possible, he suggests, is that parasites which can survive only in their host plants (obligate parasites) 'flourish best on the best fed crops.' Examples are the virus diseases, and the rust fungi of cereals. But with other parasites, including soil fungi, no generalization can be attempted. Among diseases of wheat, for example, 'take-all' is serious only on poor soils, whereas 'eyespot' is most prevalent on good land. It is the latter which provides the paradox already quoted that both parasite and individual plant are favoured by improved nutrition, with the result that there is luxuriant foliage, accompanied by a more rapid spread of infection than on poor soil.

REFERENCES

Report, East Malling Station for 1949 (1950), pp. 40, 49 and 178.
Keyworth, W. G. *Nature*, **166**, 955. (1950)
Earlier references: Bawden, F. C. *Rothamsted Experimental Station, Report for* 1939–1945, p. 131. *Plant Diseases*. London, Nelson, 1948.

RESEARCH REPORT

A. W. HASLETT

VIRUS VARIATIONS

ALTHOUGH there is good evidence that the particles of animal viruses are incapable of multiplication independently of the host-cells which they invade, the manner of their multiplication has remained speculative. Formally, the distinction between viruses and bacteria is on the basis of size: they can pass through filters, through which bacteria cannot pass. But associated with this is the further distinction that, whereas bacteria can be cultured on cell-free material in the laboratory, the multiplication of virus particles is dependent on the host-organism. The inference is that they lack the enzyme systems needed for their multiplication, and that they multiply by using the enzyme systems of their hosts.

Beyond this, there have been two principal lines of argument. One depends on the fact that there is sometimes an interval after infection during which no virus can be demonstrated. Dr F. K. Sanders, of Oxford, has shown that there is such an effect in the case of a mouse virus which in its behaviour resembles that of the poliomyelitis virus in monkeys. He has described experiments in which the virus was inoculated into the tongue muscle, and the extent of infective activity noted which was obtained from the nearest group of nerve cells serving the tongue-muscles inoculated. The virus was first detectable at about 20 hours after inoculation, this initial delay being attributed to the comparatively slow rate of travel of the virus particles (0·5 mm. per hour). The interest, in the present connexion, is in the following period. Between 30 and 60 hours, virtually no virus could be demonstrated. Yet from 60 hours onwards, there was a rapid increase, and at 100 hours, more than one million times as much virus was recovered as had originally been inoculated. The suggestion was that, during the interval of non-detection and multiplication, the virus particles had been 'disassembled'.

A second argument is based on analogy with bacterial viruses. It was shown by M. Delbrück and W. T. Bailey, in 1939, that the addition of two different strains of bacterial virus to the same culture of bacteria may lead to the appearance of two intermediate types of virus in which the characteristics of the original two strains had been differently combined. More lately, A. Novick and Professor L. Szilard, of the Institute of Radiobiology and Biophysics, the University of Chicago, have given a preliminary account of similar experiments in which a virus-strain was produced which showed the properties of one original strain in the first place, but after multiplication, those of the other. Without further implications as to mechanism, it can be said that this result, like its predecessor, provides rather strong evidence for the idea that virus particles are disassembled during multiplication. Finally, from the 'killing' or 'alteration' of virus particles by ultra-violet radiation, it has been estimated by S. E. Luria of Indiana University that an individual particle of a bacterial virus contains some 30-50 sub-units. The interpretation is that, although a single 'altered' particle is incapable of multiplying, the damaged unit in one can be made good by interchange with other particles, also altered but in different sub-units.

Evidence for the production of new strains of animal viruses, with the properties of the original strains combined in new ways, has now been described by Professor F. M. Burnet and Dr P. E. Lind, of the Walter and Eliza Hall Institute, Melbourne, Australia. Their experiments have been on the interaction, through mixed infections, between different strains of influenza virus A. One strain, common to all experiments, was first obtained by Professor C. H. Stuart-Harris, now of Sheffield University. It causes encephalitis in mice when inoculated into their brains. Like others of the strains used, this is a variant of the original influenza virus A, isolated by Sir Patrick Laidlaw and others at the National Institute for Medical Research, London, in 1934, through transmission to ferrets, and later mice. Others of the strains used have also specific properties. The key to two of these is an effect which was discovered by Dr G. K. Hirst at the Rockefeller Institute in 1941: namely, that the red blood cells of fowls

are agglutinated, or stuck together, by the influenza virus. This is because the virus particles become attached to the cell surface. One strain of the virus is distinguished by the firmness of this attachment. Another is identified by the fact that agglutination is resistant to heat. A more general distinction depends on the extent of reaction shown by antiserum prepared against one strain of the virus, with virus preparations of a different strain. By such methods, alone or in combination, the strain which produces brain-symptoms in mice could be distinguished from each of five other strains.

The procedure was to inject the brains of mice with a mixture of the two strains, the proportions of the mixture being sometimes critical. The general result was the production of strains which retained, in varying degree, the property of causing brain-symptoms in mice, but which in other respects showed the properties of the second strain. For example, if the second strain was the one which in the agglutination of fowl red cells is resistant to heat, then the virus obtained from the brains of mice after a mixed inoculation was also resistant to heat. In another case, there were 'strong indications' of the formation of a graded series of strains which differed in virulence when inoculated into mouse-brain. The possibility that the resulting virus might have been merely a mixture was excluded by further tests. One such was described by Professor Burnett in the Croonian Lecture of the Royal Society which he gave last year in London. Working with seven strains of related viruses, he has shown that these can be arranged in a constant series, according to whether red cells which had been first agglutinated by one virus could be agglutinated by another after removal of the first. From such experiments, and from others, he argues that there had been some real combination of virus strains: for there had been changes in properties.

Professor Burnett and Dr Lind conclude that their findings 'add strong circumstantial evidence in favour of the hypothesis that influenza virus particles are composite, and on entry into the host cell break down into sub-units which replicate independently, giving rise to a "pool" of virus material from which infective units can be reconstituted.' Such a mechanism may be

connected also with the changes in the relative virulence of a virus to different species, according to the host species in which it is maintained.

Brit. Med. J., ii, 675 (1950).
Novock, A. and Szilard, L. *Science*, 113, 34 (1951).
Burnet, F. M. *J. Gen. Microbiol.*, 5, 46 and 54 (1951).
Burnet, F. M. and Lind, P. E. *J. Gen. Microbiol.*, 5, 59 and 67 (1951).

AMINO-ACID SYNTHESIS

Since only 10 out of the 19 amino-acids which are incorporated in proteins are necessary constituents in the diet of young rats, and eight possibly nine for adult human beings, it has been an open question whether those not required could be built up for themselves by typical mammals, or whether they were supplied, instead, by intestinal bacteria – as with biotin, folic acid, and others of the B group vitamins. Experiments which suggest that it is probably the mammal which does the synthesis have been described lately by Dr W. C. Rose, of the Division of Biochemistry, the University of Illinois, by whom much of the most careful work on amino-acid requirements has been undertaken.

The contribution made by intestinal bacteria to the supply of B group vitamins has been shown earlier by experiments in which human subjects were given diets deficient in one such vitamin, with or without a slow-acting sulpha drug which would reduce the bacterial population. A similar procedure has been followed by Dr Rose in his amino-acid experiments, carried out on male weanling rats during a 28-day period. Three standard diets were used, with and without the sulpha drug in each case. One included all 19 amino-acids; the second, the 10 essential for rats in sufficient quantity to meet all nitrogen requirements; and the third, the same 10 amino-acids, but in individually minimum quantities, and with the balance of extra nitrogen as an ammonium salt. There were 13-22 animals in each group, and their average rate of growth was used as the index of adequacy of the diet.

In rats given the second of these diets there was a slight but significant decrease in growth in the group which was given the sulpha drug, which should imply that some small contribution was made in this case by the bacteria. But as between the two groups in which the balance of extra nitrogen was supplied as an ammonium salt, there was no significant difference depending on whether the sulpha drug was administered. Other micro-organisms, it is pointed out, which are not affected by sulpha drugs, might take part in the making of 'non-essential' amino-acids; but for the two most probable groups of bacteria, the evidence is 'rather convincing' that they do not, although for a final answer it may be necessary to experiment with germ-free mammals, as illustrated in *Science News* 20.

Rose, W. C. and Smith, L. C. *J. Biol. Chem.*, 187, 687 (1950).

BODY-WEIGHT AND FAT

Methods by which the body-weights of living human beings can be broken down into the three components of cell-mass, fat, and fluids other than in cells, have been described lately by Professor R. A. McCance and Dr E. M. Widdowson, of the Medical Research Council Department of Experimental Medicine, the University of Cambridge. They are of interest for all purposes for which the combination of weight, height, age, and apparent fitness have proved inadequate criteria. Among these are the assessment of 'ageing' and of the effects of athletic training, the progress of rehabilitation after malnutrition, and the adequacy in kind as well as quantity of particular diets. The need for such methods can be illustrated both by past research by the same two workers, and by a recent study of 'ageing' from the University of Minnesota.

Taking the Cambridge experience first, Dr Widdowson reported in 1948 and 1949 on successive stages in the rehabilitation of German orphanage children after the war, with results which presented some awkward problems to nutritionists. Eating normal German rations, with only half a pint to a pint of milk per week, but with the addition of unlimited bread which for different groups of children was of different degrees of extraction, these children made nearly two years' progress in weight and height in a single year, and at the end of that time came close to American standards for their ages. Nor was there any difference according to the quality of the bread which they had been given. It would have been useful in such an experiment to be able to distinguish between fat, cell-mass, and fluid, both at different stages of recovery and compared with 'normal' children.

The Minnesota research on 'ageing' has been described lately by Dr J. Brozek and Prof. Ancel Keys, and suggests the same point. Their comparison was between two groups of men, aged 19-22 years and 45-55 years respectively, but almost identical on the average in the relation of body-weight to height. It was concluded that, in the older group, there had been a replacement of muscle and other active tissues by fat, but the evidence for this view consisted in a 2 per cent reduction in specific gravity and an increase in wrinkling. Again, a comparison of fat and cell-mass would have been preferable.

The method used at Cambridge depends on two independent measurements of water – one of total body water, and the other of water not contained in cells. The difference gives the water-content of cells; and from this the cell-mass is calculated, on the assumption that 67 per cent of the cell-mass is contributed by water. The originality is rather in the use made of these measurements than in the measurements as such. For the measurement of total body-water, the requirement is for a substance which can be readily administered, and distributes itself in a uniform way through all the water in the body. Urea taken by mouth is the standard method, and this has been used. For the measurement of water other than in cells, the requirement is for a substance which does not enter cells, but is uniformly distributed otherwise. The nearest approach is thiocyanate, given intravenously. This enters the red cells, but not others, so that only a relatively small correction is needed to obtain the amount of extra-cellular water. The use of thiocyanate, also, is now a standard method. Given these two figures, the amount of water in cells is obtained as a difference, as already stated, and from this the cell-mass is calculated. Finally, the weight of fat in the body is obtained as the balance which is left over after cell-mass and extracellular fluid have been subtracted from body-weight. For seven 'normal'

men, the figures ranged from 7·2 per cent to 22·4 per cent fat, and for individuals described as 'obese' from 36·7 per cent to 44 per cent. Incidentally, the load of fat carried by the biggest of the normal men amounted to about 40 pounds which 'would be a big one for a rucksac'. The figures for normal women overlapped with those for normal men, but on the average were higher, and that for the fattest woman was 58·1 per cent. It is added, a little unkindly, that such a proportion could only be attained in livestock by special breeding.

As an example of use, measurements of both kinds have been made on a group of 10 under-nourished German men during rehabilitation in hospital. It is recorded that they 'ate very large quantities of food, but that little or nothing was done for them in the way of graded exercises.' When they came to hospital, they averaged 127 pounds weight, and had a smaller cell-mass than usual, not only absolutely but in proportion to body-weight. Fat-content was also low, though in some individuals this was within normal limits. On the other hand, the proportion of water not in cells was higher than usual. Two months later, they had put on an average of 15·5 kg. weight. But the amount of fat which they had acquired was less in weight than would have been suggested by appearance; and in the proportion both of cell-mass to fluid, and of fat to fluid, they were still abnormal. Other intended uses of the method are in the study of development during childhood and adolescence, in relation to the changes which take place during athletic training, and for comparisons between the various indices of nutritional state which 'have been such a disappointment to many of us.'

McCance, R. A. and Widdowson, E. M. *Proc. Roy. Soc.* 'B', **138**, 115 (1951).

Brozek, J. and Keys, A. *Science*, **112**, 788 (1950).

RADIO ASTRONOMY

A postscript can already be added to the article by A. J. Higgs on the same subject which is printed earlier in this issue. It concerns the radio 'outbursts' which are characterized as being of relatively long duration and great intensity compared with the shorter 'bursts' to which he also refers. The most intense outbursts are typically associated with visual solar flares, beginning at the shorter radio-wavelengths and progressing to longer wavelengths, up to minutes later, the change in wavelength being attributed to the outward movement through the sun's atmosphere of the disturbance by which they are caused. In the interval, an unusually interesting case has been described by G. Reber of the U.S. National Bureau of Standards, Washington, from observations made on nine fixed frequencies by workers at the Bureau of Standards, the U.S. Naval Research Laboratory and at Cornell University. The interest is in a departure from the usual sequence of frequency change towards the later stages of the outburst. This was on July 12 of last year, and on the highest frequency recorded (9,500 Mc./s) began at 1107 hours EST. From then on, there was the usual shift towards lower frequencies, with increased emission beginning at 1116 hours at 51 Mc./s. This continued for about 4 minutes, compared with durations of from 5 to 7 minutes at higher frequencies. It was at this point that there was a departure from the usual, as generalized by Higgs from Australian observations. At the next lower frequency, activity continued longer: this was at 35 Mc./s (equivalent wavelength about 86 metres). Compared with the previous comparatively short durations of increased emission, that at 35 Mc./s continued without break for 25 minutes; and there was no observable increase at the next lower frequency. Instead, there was a second period of increased emission on the previous frequency (51 Mc./s), this being separated from the first period of activity by an interval of about 3 minutes. Also, the increased emission built up gradually, instead of quickly as in other cases.

The interpretation offered by Reber is described as frankly speculative. His suggestion is that the material causing the disturbance had reached the top of its trajectory at a level close to that corresponding with the generation of waves of 35 Mc./s frequency; and that the second period of emission at 51 Mc./s was connected with the return of erupted material, now falling back towards the sun's surface, but becoming already quite

'weak and dispersed', with the result that at lower levels (the next corresponding with 75 Mc./s emission) there was no resumption of activity.

There is an obvious interest in relation to the delayed effects, which on some occasions, but not always, are produced in the atmosphere of the earth, and which for some time have been connected with particle-emission from the sun – as distinct from the immediate effects of a solar flare, which are caused by an increase in ultra-violet radiation from the sun at approximately the same time as the visual flare. Among the delayed effects are auroral displays, the interruption of radio communications on routes crossing the polar regions, and major magnetic disturbances. They are found only after flares seen on those regions of the sun's disc which face in the general direction of the earth, and then not on all occasions. This broadly directional requirement, together with the delay in time from the visual flare, was part of the earlier evidence that these delayed effects were caused by particles from the sun. From their concentration on polar regions, it was further concluded the particles were electrically charged; and, since a visual solar flare can be seen with a spectroheliograph to be a bright hydrogen eruption, it was natural to expect that the charged particles would prove to be protons or hydrogen nuclei. It has been confirmed lately by workers at the Yerkes Observatory, the University of Chicago, and at Cornell University that aurorae are, in fact, accompanied by a stream of incoming protons, their apparent motion being consistent with the spiral path towards the earth's magnetic poles that must be followed by charged particles approaching the earth's atmosphere from outside.

The relevance of the Bureau of Standards' observations is to the 'sometimes but not always' quality of these delayed effects. The picture given by the Australian radio observations fits well with the case in which visual flare, radio outburst and delayed effects are associated as consequences of the same initial disturbance. In the case of the outburst described by Reber, there was only a probable association with a visual flare, not itself outstanding. The final suggestion is that 'in the particular phenomenon observed, the eruption came just short of leaving the sun' and so was incapable of creating magnetic or other disturbances on the earth as delayed effects, irrespective of the directional factor. Obviously, no general conclusion can be drawn from a single such outburst. What seems significant for solar physics is that the connexion, if real, is of a kind which can be tested.

Reber, G. *Science*, **113**, 312 (1951).
Meinel, A. B. *Science*, **112**, 590 (1950).
Gartlein, C. W. *Phys. Rev.*, **81**, 463 (1951).

OCEANOGRAPHY

Two examples of the use of new methods of research in oceanography may be brought together. The first, giving an easy measure of the rate of carbon fixation by marine algae through photosynthesis, is reported by Professor E. S. Nielsen, of the Botanical Department, the Royal Danish School of Pharmacy, Copenhagen, from experiments carried out in the eastern Atlantic between Lisbon and Cape Town during the Danish oceanographical expedition in *Galathea*. Measurements were made by a tracer method on samples of ocean-water taken in 58 positions and at different depths. The results confirmed the expectation that 'by far the greatest production of organic matter on earth is due to marine plankton algae.'

The method depends on the rate of absorption by algae of bicarbonate solution, labelled with carbon-14, which is radioactive. Small, measured quantities of radioactive solution were added to specimens of sea-water containing its natural population of algae. In the most direct version of the method, these were then exposed to natural illumination in sea water. This was done between noon and sunset with the ship stationary, and presumably at the same depths as that from which the original specimens had been collected. Radioactive carbon was assumed to be assimilated at approximately the same rate as normal carbon; and, the algae being then filtered off, and inorganic carbon removed from them by acid vapour, their observed radioactivity gave a measure of the total carbon assimilated. In other experiments, the temperature and intensity of illumination were

recorded at different depths with the ship in motion, and the specimens of sea-water collected exposed to experimental illumination on shipboard at the same temperature. The results were compared with those given by the more direct method when the ship was stationary. Possible complications from the exchange of radioactive for stable carbon, as between sea-water and algae, are omitted from consideration in the preliminary report, though it is recorded that preparatory tests were made in Copenhagen, and there would have been a further check from the making of observations at different depths and so at different intensities of illumination.

With this qualification, the rate of carbon-fixation was found to vary between about 20 mgm per square metre of ocean surface per day near Teneriffe, to about 2,000 mgm in the neighbourhood of the Berguela current, where there is an upwelling of deep ocean water, rich in nutrient salts. Carbon-fixation at the rate of 100–200 mgm per square metre per day was most commonly found. For comparison, the rate of timber-production per square mile of forest has been quoted as averaging between 15,000 and 20,000 cubic feet of timber per annum. Taking the carbon content of wood as 25 per cent, the former figure corresponds with a rate of carbon-fixation of the order of 100 mgm per square metre per day, averaged over the year. It would seem therefore that the rate of carbon-fixation is as great or greater in the oceans as on land, in proportion to surface area; and since 70 per cent of the earth's surface is covered by ocean, it follows that 'by far the greatest production of organic matter' is in ocean water.

The second innovation is the application to ocean-bed research of the method of seismic prospecting which has given most information when used on land. The principle is to detonate an explosive charge at a known depth and to measure the time of arrival of the resulting shock-waves at a series of points which are spaced out in a straight line from the point of explosion. In general, the rate of propagation is greater in layers which are at a greater depth beneath the surface; and in favourable conditions, both the depth and the thickness of successive layers can be worked out. The procedure is known as 'refraction shooting', as distinct from 'reflexion shooting', which is a straightforward echo method.

The application of the refraction method in deep ocean water was undertaken by M. N. Hill and J. C. Swallow, of the Department of Geodesy and Geophysics of Cambridge University. Preliminary experiments were carried out south of the Eddystone lighthouse in 1947, with shipping and other facilities provided by the Marine Biological Association, Plymouth. Further work was based on a weather ship, stationed some 350 miles off the west coast of Eire in the position, 53° 50′ N., 18° 40′ W. Depth-charges, set to fire at 900 ft. depth, provided the shock waves; and recording was by quartz hydrophones suspended at 150 ft. depth from a series of radio-transmitting buoys. The practicability of the method being thus established, five further sets of observations have been made by T. F. Gaskell and J. C. Swallow during a survey cruise in the North Atlantic by H.M.S. *Challenger*. Including the weather-ship measurements, there are thus six sets of observations. Of these, two are in the eastern Atlantic, one in the central Atlantic on the western side of the Atlantic Ridge (41° 46′ N., 35° 30′ W.), and three more in the neighbourhood of the Bermudas.

The interest is in the thickness of sediments found in different positions, and in the apparent variation in the thickness of the granitic layer, compared with that under the continents. In the two eastern positions, the thickness of sediments were estimated at around 3,000 ft and 7,700 ft respectively; in the central Atlantic at only 800–1,200 ft; and in the positions near the Bermudas at 2,400 ft to 14,000 ft. In the weather-ship position, the thickness of the granitic layer was estimated at 9,000 ft, or about one-third to one-fifth of that beneath the continents. On the other hand, there was no indication of a granitic layer in the central Atlantic; and, with one doubtful exception, it was again missing in the records obtained at the three remaining stations near the Bermudas.

Nielsen, E. S. *Nature*, **167**, 684 (1951).
Hill, M. N. and Swallow, J. C. *Nature*, **165**, 193 (1950).
Gaskell, T. F. and Swallow, J. C. *Nature*, **167**, 723 (1951).

CORRESPONDENCE

SYNTHETIC FIBRES

MY friend, Professor Wyckoff, has sent me *Science News* 14, which I have read with great pleasure. I note, however, that some data in the article contributed by Mr Moncrieff are not correct. I give below a corrected version of these facts, about which I am sure that Mr Moncrieff will be glad to be informed.

(1) He ascribed the 'outstanding discovery' that the molecules of fibres are very long to Astbury. This is not true: many years before Astbury's publications in this field, the threadlike nature of fibre-forming molecules had been proved by Staudinger and Haworth, and the first conclusions about the arrangement of these long molecules in cellulose (based on *x*-ray evidence) were due to the American, Sponsler[1]. The atomic arrangement in cellulose, i.e. the exact position of the chain molecules, also based on *x*-ray diagrams, was established by Meyer and Mark.[2] The same authors were also the first to establish the existence of fibrous proteins and the presence of extended protein chains in silk, tendon, muscle, etc.[3]

(2) The first synthetic fibre was not nylon, but PeCe, produced and marketed by I.G. as early as 1934, under the name of W. K. Silk.

KURT H. MEYER

1. *Colloid Symposium Monograph* 4, p. 174 (1926).
2. *Der Aufbau der hochpolymeren organischen Naturstoffe*. Leipzig, 1930.
3. *Berichte der deutschen chm. Ges.* **61**, 1932 (1928).

Mr. R. W. Moncrieff writes:

I have noted Professor Meyer's remarks, and hope that I may have the opportunity of replying to them as follows:

(1) It is, of course, perfectly true that Mark, Meyer, Herzog, Sponsler, Hess, Howarth, Staudinger and many others had worked on the structure of cellulose and fibroin, and had

established the linearity of their molecules before Astbury's work on the keratins was published. Nevertheless, I still think it is true to say that it was Astbury who first appreciated the 'fibrous' nature of the molecules of fibres: to quote from his paper (with Street) in *Phil. Trans. Roy. Soc.*, London, 1931, **230** A, 96:

> The combined evidence of chemical and *x*-ray analysis, when brought to bear on the problem of fibre structure in both organic and inorganic fields, is undoubtedly leading to a relatively simple generalization; in fact, to a conclusion no more profound than that fibres are what they are because their inner molecular structure is also of a fibrous nature.

Later, in his admirably lucid book *Fundamentals of Fibre Structure*, Astbury forced this conclusion on the attention of a wide range of scientists and technologists.

In scientific work, it is only seldom that the man who first breaks new ground is the one to harvest the crop. The Periodic Classification is universally attributed to Mendeléeff, yet the fundamental discovery of periodicity was made by Newlands; the Atomic Theory is unblushingly ascribed to Dalton, although there were a dozen before him who had dealt in atoms. Similarly, to have ascribed our knowledge of 'fibrous' molecules to Astbury was, I believe, not unfair, particularly in a highly condensed introduction to an article intended not primarily for fibre specialists but for scientists generally.

(2) In reply to Professor Meyer's second point, I must confess that I have never come across any reference to 'W. K. Silk'.

BOOKS RECEIVED

The Origin of the Earth by W. M. Smart (Cambridge University Press, 1951), pp. 239. 12s. 6d.

Professor Smart discusses the problem of the origin of the earth at a level comparable with that of the more general articles in *Science News*. In the first part of his book, he describes the solar system, including the earth, and ends with the conclusion that there can be no case for the 'piecemeal growth of the solar family': the uniformities are too many. In Part II, 'When?', he gives the evidence on chronology, with reasonable agreement up to this stage. In Part III, 'How?', theories of the origin of the solar system are briefly summarized and discussed, beginning with the Kant-Laplace theory of aggregation from a nebula, and ending with a supernova-explosion of a former companion-star of the sun, as postulated by Hoyle. The number of recent theories is quoted as presumptive evidence that 'nothing approaching finality ... has so far been reached.'

The History of Nature by C. F. Von Weizsacher (Routledge and Kegan Paul, 1951), pp. 180. 12s. 6d.

This is a patently sincere book by a theoretical astronomer, by training a physicist, who feels that 'specialized science is powerless to give us a world view that could sustain us in the confusion of our existence', and believes that choices in the use of knowledge 'do not lie solely in the sphere of morals, but in that of religion'.

Quantum Mechanics by A. Landé (Pitman, 1951), pp. 307, 40s.

The level is that of the university student. The author, who is Professor of Physics in the Ohio State University, gives an introduction to quantum mechanics in which the emphasis is on physical background and historical development rather than on a rigid and complete formulation as a mathematical method. The form is frankly didactic in places, and assumptions and results are not always distinguished.

ABOUT OUR CONTRIBUTORS

H. J. BHABHA, F.R.S., is Director of the Tata Institute of Fundamental Research, Bombay, and Professor of Theoretical Physics at the Institute. He is at present Chairman of the Indian Atomic Energy Commission. Aged 41, he worked with P. A. M. Dirac at Cambridge. He is, with W. Heitler, the author of the cascade theory of cosmic radiation mentioned in his article.

W. K. BURTON graduated in electrical engineering at Manchester in 1943 and, after proceeding to a higher degree in engineering joined the staff of I.C.I. Ltd. as a theoretical physicist. Since 1947, he has been seconded to the H. H. Wills Physical Laboratory, University of Bristol, from I.C.I. Ltd., Butterwick Research Laboratories. He took a Ph.D. in theoretical physics at Bristol in 1950.

J. J. DE GRYSE was born and educated in Belgium. As an entomologist, he was early interested in morphology. He has worked in forest entomology for some 35 years, and at present is Chief of the Division of Forest Biology, Science Service, Department of Agriculture, Canada.

JOHN HANCOCK graduated at the University of Helsingfors in 1935. He has since been employed in practical, administrative and research activities in agriculture in Finland, Australia, New Zealand and the South Sea Islands. He is now Chief Animal Geneticist, Ruakura Animal Research Station, Department of Agriculture, Hamilton, New Zealand. Interests: Animal breeding and behaviour.

A. J. HIGGS was formerly on the staff of the Commonwealth Observatory, Canberra, and interested primarily in ionospheric research. He joined the Commonwealth Scientific and Industrial Research Organization to undertake radar development early in the Second World War, and since 1945 has been Technical Secretary of the Organization's Division of Radiophysics.

G. C. MCVITTIE is well known as an exponent of the observational approach to the cosmological problem. During the Second World War, he was seconded for special service with the Meteorological Office, and for his work as scientist and administrator was awarded the O.B.E. In 1948, he was appointed to the newly created chair of mathematics at Queen Mary College, University of London.

SCIENCE NEWS

Outstanding articles from recent issues include:

SCIENCE NEWS 17

Physics and Metaphysics	*Prof. Max Born*, F.R.S.
My Attitude to Quantum Theory	*Prof. Albert Einstein*, FOR.MEM.R.S.
Finding Water	*Sir Cyril S. Fox*
The Instrumentation of the Living Body	*J. W. S. Pringle*
The Early History of Man	*L. S. B. Leakey*
The Shower Cloud	*James Paton*
Semiconductors	*James T. Kendall*

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 To-day	*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*
Controlled Corrosion	*F. A. Fox*
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*
The Science of Lubricating Greases	*Thomas D. Smith*
A Definition of Man	*Kenneth Oakley*
The Origin of Language	*R. A. S. Paget*