Physics

Physics and the engineer

Nuffield Advanced Science

Copy 2

Physics and the engineer

082723 4
anced Science

A

Nuffield Advanced Physics team

Joint organizers

Dr P. J. Black, Reader in Crystal Physics, University of Birmingham
Jon Ogborn, Worcester College of Education; formerly of Roan School, London SE3

Team members

W. Bolton, formerly of High Wycombe College of Technology and Art
R. W. Fairbrother, Centre for Science Education, Chelsea College; formerly of Hinckley Grammar School
G. E. Foxcroft, Rugby School
Martin Harrap, formerly of Whitgift School, Croydon
Dr John Harris, Centre for Science Education, Chelsea College; formerly of Harvard Project Physics
Dr A. L. Mansell, Centre for Science Education, Chelsea College; formerly of Hatfield College of Technology
A. W. Trotter, North London Science Centre; formerly of Elliott School, Putney

Evaluation

P. R. Lawton, Garnett College, London

Cover illustration
This shows a modern miniaturized computer component. It is an integrated circuit, etched and deposited on a single chip of semiconductor material. Each chip, though containing the equivalent of many transistors, resistors, etc., is only about 1 mm across, and contains all the circuits needed for a random access memory unit capable of storing 1024 bits of information.
Photograph, Mullard Ltd.

Physics and the engineer

Nuffield Advanced Science
Published for the Nuffield Foundation by Penguin books

Penguin Books Ltd, Harmondsworth, Middlesex, England
Penguin Books Inc., 7110 Ambassador Road,
Baltimore, Md 21207, U.S.A.
Penguin Books Ltd, Ringwood, Victoria, Australia

Design and art direction by Ivan and Robin Dodd
Illustrations designed and produced by Penguin Education

Filmset in 9 on 11 pt 'Monophoto' Univers Medium
by Keyspools Limited, Golborne, Lancashire
and made and printed in Great Britain
by C. Tinling & Co. Ltd, Prescot and London

Contents

Consultative Committee

Professor Sir Nevill Mott, F.R.S. (Chairman)
Professor Sir Ronald Nyholm, F.R.S. (Vice-Chairman)
Professor J. T. Allanson
Dr P. J. Black
N. Booth, H.M.I.
Dr C. C. Butler, F.R.S.
Professor E. H. Coulson
D. C. Firth
Dr J. R. Garrood
Dr A. D. C. Grassie
Professor H. F. Halliwell
Miss S. J. Hill
Professor K. W. Keohane
Miss D. M. Kett
Professor J. Lewis
J. L. Lewis
A. J. Mee
Professor D. J. Millen
J. M. Ogborn
E. Shire
Dr J. E. Spice
Dr P. Sykes
E. W. Tapper
C. L. Williams, H.M.I.

Foreword

It is almost a decade since the Trustees of the Nuffield Foundation decided to sponsor curriculum development programmes in science. Over the past few years a succession of materials and aids appropriate to teaching and learning over a wide variety of age and ability ranges has been published. We hope that they may have made a small contribution to the renewal of the science curriculum which is currently so evident in the schools.

The strength of the development has unquestionably lain in the most valued part that has been played in the work by practising teachers and the guidance and help that has been received from the consultative committees to each Project.

The stage has now been reached for the publication of materials suitable for Advanced courses in the sciences. In many ways the task has been a more difficult one to accomplish. The sixth form has received more than its fair share of study in recent years and there is now an increasing acceptance that an attempt should be made to preserve breadth in studies in the 16–19 year age range. This is no easy task in a system which by virtue of its pattern of tertiary education requires standards for the sixth form which in many other countries might well be found in first year university courses.

Advanced courses are therefore at once both a difficult and an interesting venture. They have been designed to be of value to teacher and student, be they in sixth forms or other forms of education in a similar age range. Furthermore, it is expected that teachers in universities, polytechnics, and colleges of education may find some of the ideas of value in their own work.

If the Advanced Physics course meets with the success and appreciation I believe it deserves, it will be in no small measure due to a very large number of people, in the team so ably led by Jon Ogborn and Dr Paul Black, in the consultative committee, and in the schools in which trials have been held. The programme could not have been brought to a successful conclusion without their help and that of the examination boards, local authorities, the universities, and the professional associations of science teachers.

Finally, the Project materials could not have reached successful publication without the expert assistance that has been received from William Anderson and his editorial staff in the Nuffield Science Publications Unit and from the editorial and production teams of Penguin Education.

K. W. Keohane
Co-ordinator of the Nuffield Foundation Science Teaching Project

Introduction

Magazines such as *Science Journal* have performed an important service in bringing to a wider public the writing of scientists and engineers, so making it possible for the scientifically informed reader to keep up with new developments in science and technology. The availability of reprints from such magazines has, for educational purposes, been a particularly useful service, providing a convenient and cheap form of up to date and interesting information for students, written by leaders in each field.

Since *Science Journal* has been discontinued as a separate magazine, reprints from it are no longer available. This volume, containing eight such reprints, has therefore been produced. The reprints included in it have been chosen for their technological interest, and for their direct usefulness in conjunction with the Nuffield Advanced Physics course. They mainly concern two areas of technology which are emphasized in the Nuffield course: materials and electromagnetic machines. The articles have been brought up to date by the authors where they thought it desirable.

It is hoped that, while these reprints will be of particular value for the Nuffield course, they will also interest a wider audience. Anyone who is interested in how engineers think, and in how physical principles can be used in the service of human needs, should find something of value in this collection.

Original publication dates

The papers collected in this book were first published, in their original form, in the following editions of *Science Journal.*

Electrostatic engineering: September 1965
Tall buildings: October 1965
High temperature materials: August 1965
New forms of electric motor: February 1966
Colour television: April 1966
Fibre-reinforced metals: November 1966
The homopolar generator: February 1967
Ultrahigh magnetic fields: May 1966

N. J. Felici

Electrostatic engineering

Although their potential is still seldom appreciated, electrostatic devices are already saving Western European industry some £10 million a year. During the next few years they could lead to a new growth industry.

'The advent of electrostatic equipment has been one of the major advances in technology in the last 30 years.'

This opinion, from a leading European industrialist, may seem puzzling as electrostatics is usually thought of as an old fashioned chapter of electrical science of little practical interest. Indeed, industrial applications had hardly any significance before the 1920s, and their rapid expansion started only 15 years ago. Even now interest in applied electrostatics is meagre compared with other technological developments such as automation and the conquest of space. But, as I shall describe, it is growing ever more rapidly.

It would be futile to try to distinguish precisely between electrostatics and conventional electrical engineering. Electrostatics is concerned primarily with the forces exerted at a distance by any electrified body – one with an excess or deficiency of electrons – in contrast to electromagnetic forces exhibited by metals of the iron family or by strong currents in good conductors. But whereas few substances respond significantly to magnetism or conduct strong currents, electrostatic forces can act on all matter. This is the key to their industrial importance but, until the advent of modern technology, they were too feeble and erratic to form the basis of a reliable everyday industrial process.

Electrostatic forces

Another characteristic of electrostatics is the occurrence of very high voltages, sometimes in excess of 100 000 volts. The factor controlling electrostatic forces is the 'electrostatic field strength' measured in terms of voltage per unit length. Two parallel metal plates 0.1 m apart with a potential difference of 40 000 V will give a uniform field strength of $4 \times 10^5\ \mathrm{V\,m^{-1}}$. This is not a very intense electric field and it gives an electrostatic pull of only $0.7\ \mathrm{N\,m^{-2}}$ – barely sufficient to lift fragments of paper. Industrial processes can therefore rarely take advantage of electrostatics except at high voltages. Originally this demanded bulky and expensive equipment with elaborate safety locks, but today a portable spray gun can be supplied with 90 000 V d.c. by a lightweight electrostatic generator without danger even if the core of the cable happens to become exposed.

The fact that field strength is the controlling factor of electrostatic forces also prevents electrostatic engineers from insulating conductors by mere separation. On the contrary they must put live conductors or electrodes as close together as possible to operate at the brink of flashover. Equipment must therefore be designed to tolerate sparks and arcing.

Moreover, electrostatic forces are dependent on surface area, whereas gravitational, inertial, and magnetic forces are volume dependent. This means that the smaller the piece of matter involved, the greater the relative strength of the electrostatic force. Thus electrostatics owes most of its importance to the fact that finely divided matter occurs throughout industry – deliberately as in painting or crop spraying, or discharged as an unwanted by-product such as fly-ash and fumes. Most natural ores, for instance, are complex conglomerates of very fine crystals which cannot be physically separated unless they are previously freed by grinding. The electrical properties of the grains often differ enough for electrostatic forces to pick up one component while rejecting the others.

If we scale down an electrified piece of matter to atomic or subatomic size we obtain ions or electrons which are so small as to be extremely sensitive to an electrostatic force. For this reason acceleration of elementary particles in vacuum by means of electrostatic fields has become very important in the study of radiation and nuclear physics. X-ray tubes themselves are a form of small electron accelerator, but this term is usually restricted to devices dealing with ions and operating at over 100 000 V; the many commercial models are mainly used for fast neutron generation. Recently they have found a novel application in space technology in which minute spheres of diamond or tungsten, about one micron in size, can be electrostatically accelerated in vacuum to speeds from 1 to 80 kilometres per second to simulate the impact of micrometeorites on space vehicles.

Electrostatics can thus serve as a kind of motive power. The converse can also be put to use; mechanical work can be turned into electrical energy, mostly in the form of high voltage d.c., by means of electrostatic forces in a rotating machine usually referred to as an 'electrostatic generator'.

Generators

Electrostatic generators are unique in their ability to produce very high voltage direct current, without intermediate step-up transformers and rectifiers, as easily as a conventional dynamo gives low voltage d.c. For this reason they have in the last 30 years enjoyed a revival after a half century of near oblivion. At low power outputs below ten kilowatts their power-to-weight ratio is good, but they cannot compete effectively with electromagnetic machines when high power is required. This is because of the relative weakness of electrostatic attraction compared with that of strongly magnetized bodies. In air at atmospheric pressure the maximum electric field that can be reached without sparking is about 3×10^6 V m^{-1}, giving an electrostatic pull of about 35 N m^{-2}, whereas a strongly magnetized pole-piece can attract a soft iron armature with a pull of 700 kN m^{-2} or more. Electrostatic generators must resort to pressure insulation to allow for stronger electric fields and higher specific power.

At present field strengths in excess of 3×10^7 V m^{-1} can be achieved in hydrogen or carbon dioxide at 20 times atmospheric pressure to improve electrostatic forces and power output by a factor of 100 as compared with free atmosphere operation. This is sufficient to arouse interest in the 10 to 10 000 watts range, at 50 000 to 10 000 000 V and current intensities from a few hundred microamperes to dozens of milliamperes.

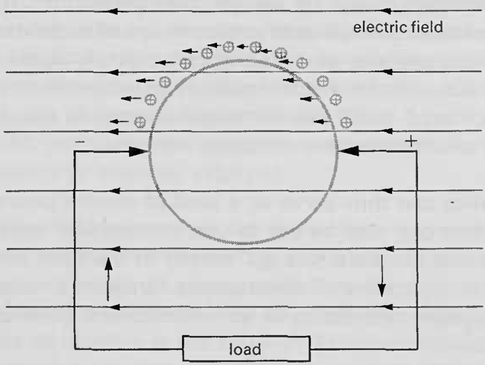


Figure 1
Simple electrostatic generation of d.c. power can be accomplished by a charge carrier (tinted ring) rotated in a stationary electric field. Positive charge is sprayed on at the negative pole on the left and carried round against the field to a higher potential for collection by the positive pole and delivery to the load. The net effect can be doubled if the lower half of the carrier is sprayed with negative charge from the positive pole. The system can also work as a motor if the charge is fed the other way.

The chief advantages of electrostatic generators are low internal capacitance, perfect smoothness, and self-limitation of current, which renders them completely safe for human beings and associated equipment. I have pointed out that electrostatic engineering cannot accept broad safety margins as far as flashover and sparking are concerned, and that 'brinkmanship' operation is almost mandatory for the sake of efficiency. Electrostatic generators are well fitted to obviate resulting hazards. No sudden release of energy is to be feared in the case of sparking – nor is the short circuit current dangerous because it is ripple free, smooth d.c., the 'let go' and 'shock' thresholds of which are much higher than those of a.c. or pulsating d.c. The same is true for the equipment, be it a paint sprayer or accelerating tube; it cannot be harmed by recurrent sparking.

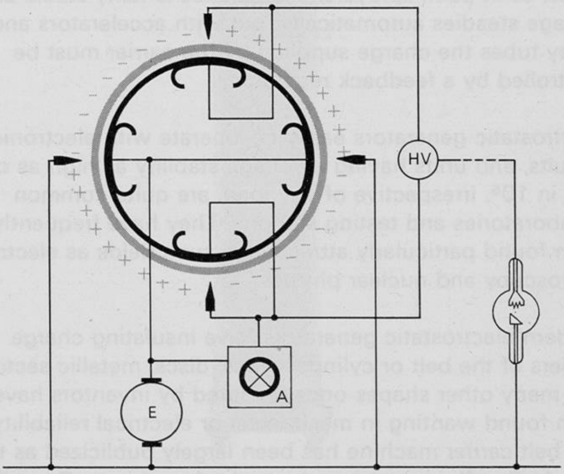


Figure 2
The four-pole electrostatic generator has a rotor (tinted) revolving around a stator inside which are four metallic conductors, two connected to the high voltage output (HV) and two to the exciting source (E) which is controlled by a feedback amplifier (A) to match current output to load requirements and keep terminal voltage constant. Charges are sprayed on to the rotor by ionizing blades, two connected to HV and two to earth; the stator is slightly conductive and smooths field distribution between adjacent ionizers. The generator delivers 0.8 to 1.6 watts per square centimetre of rotor and 12 to 16 kilovolts per centimetre spacing between ionizers. Generators of this kind are frequently used to power X-ray tubes (right).

The heart of an electrostatic generator is an endless charge carrier which transports a charge from an earthed electrode to an insulated high voltage terminal which is thus fed with a controllable current. Were it not for charge leakage, corona discharge, or conduction, the potential could surge to enormous values in a very short time. In fact the voltage of an electrostatic generator is limited only by the insulation of the terminal. Since there is no definite electromotive force in an electrostatic generator the actual voltage depends on a varying balance between charge leakage, current drain from external load, and current supply from the carrier. With some loads, as in paint sprayers, this balance is fairly stable and the voltage steadies automatically; but with accelerators and X-ray tubes the charge supplied by the carrier must be controlled by a feedback regulator.

Electrostatic generators easily co-operate with electronic circuits, and units having a voltage stability as high as one part in 10^5, irrespective of any load, are quite common in laboratories and testing stations. They have frequently been found particularly attractive in such fields as electron microscopy and nuclear physics.

Modern electrostatic generators have insulating charge carriers of the belt or cylinder types; discs, metallic sectors, and many other shapes once favoured by inventors have all been found wanting in mechanical or electrical reliability. The belt carrier machine has been largely publicized as the Van de Graaff generator and is to be found in nuclear laboratories all over the world. It is highly suitable for voltages in the range 1–10 million volts but rarely gives currents in excess of one milliampere. The hydrogen insulated cylinder machine differs greatly in structure and performance although it operates on similar physical principles. Its voltage range is only 50 000 to 1 000 000 V, but the current output of 0.3 to 50 milliamperes is to be considered substantial by electrostatic standards.

Such a voltage and power range meets most of the current needs of industry for low power, very high voltage d.c. Stationary and portable paint and powder sprayers, X-ray therapy equipment, cable and alternator test gear, electron microscopes, and portable accelerators have all used cylinder generators as power supplies, with the inherent advantages of safety, efficiency, and compactness.

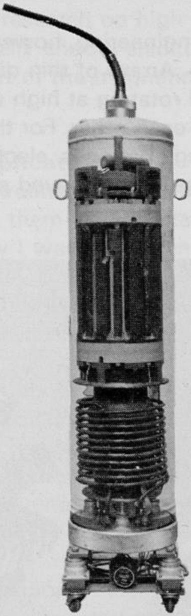


Figure 3
The eight-pole electrostatic generator of the cylindrical type is housed in a steel pressure tank (superimposed picture), about 1.5 m high, to enable field strength to be increased. Below is the 2.6 kW three phase a.c. motor and cooling coil; at the top is the 140 000 V, 2000 W, d.c. generator and output. Conversion efficiency is 70 to 75 per cent.

The feasibility of much more powerful electrostatic generators, which might solve the problem of high voltage d.c. power transmission in a straightforward way, has been the subject of considerable speculation. Taking into account the field intensities available in good fluid insulants with low frictional losses, such as compressed gases, power-to-volume ratios comparable with those achieved by electromagnetic machines would be reached by electrostatic generators with interleaved multiple discs. Such a design would make full use of its volume by having multiple live parts which could compensate for the lower specific forces afforded by electrostatic fields.

From the standpoint of engineering, however, such proposals seem rather questionable. Arrays of thin discs or sectors, alternately stationary and rotating at high speed, are likely to vibrate, rub, or even foul each other. For these reasons, unlike capacitors or storage batteries, electrostatic generators cannot take full advantage of interleaved parts.

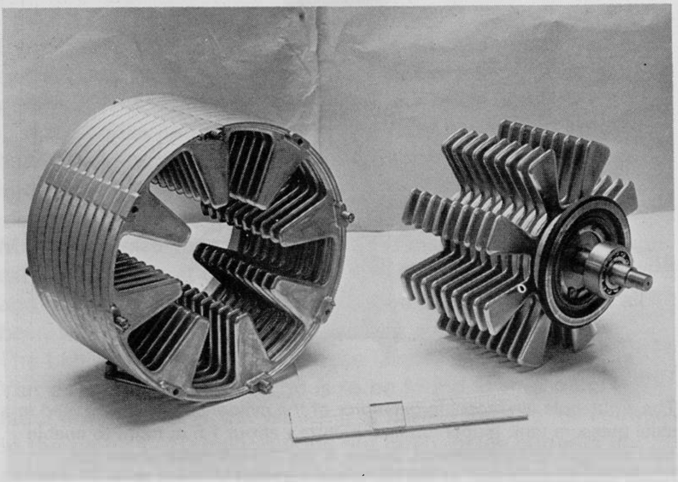


Figure 4
Stator and rotor of a variable capacitance electrostatic generator. Vacuum-insulated machines of this type are being evaluated as a source of high voltage power for spacecraft.

At any rate, the most realistic approach is to dispense with insulants in the rotor assembly altogether by having something like a rotating capacitor on a shaft supported by stationary insulators. Such a machine, worked in high vacuum, has been proposed as a possible voltage source for spacecraft ion rockets, since it is the only type of generator capable of working indefinitely under extremes of temperature and radiation. As far as terrestrial power units are concerned, perhaps a more successful line of attack may spring from present research on high permittivity liquids. These might eventually endow electrostatic engineering with the exact counterpart of the magnetic cores of conventional machines.

Industrial applications

Applications of electrostatics are so numerous that it would be impossible to list them all. For example, in a small Japanese tea factory I was shown an ingenious separator that picks up the dried leaves and rejects the worthless stems and wooden fragments. It was engineered locally and has never, to my knowledge, been mentioned in any review.

Electrostatics is the only universal way of directly using electrical energy for the displacement, acceleration, or sorting of matter. Electricity has been otherwise merely a means of transmitting energy from a remote prime mover, needing clumsy and inefficient energy conversion at the receiving end of the link.

The first major application of electrostatics was the precipitation of dusts, fly-ash, and fumes. Electrostatic precipitation now plays an outstanding role in reducing air pollution. Every modern thermal power station has huge precipitators which eliminate 95 to 99 per cent of the fly-ash discharged to the stack. Pulverized fuel firing, which made low grade coal suitable for large power plants, also made electrostatic precipitation necessary because the unburnable 20 to 40 per cent of the fuel is converted entirely to fly-ash which escapes with the flue gases. A 250 megawatt station produces about 30 tonnes of ash per hour which, distributed over one square kilometre around the station, would in a few years result in a layer several centimetres thick. A reliable electrostatic precipitator is therefore indispensable.

A modern precipitator is robust and simple, requiring no maintenance other than periodic cleaning with hot water sprays. It consists of a matrix of metal channels, each about 0.30 m wide and about 1 m high, containing arrays of insulated steel wires maintained at a negative d.c. voltage of 30 000 to 60 000 V. Corona discharge, sometimes assisted by barbs or wedges on the wires, creates a flow of negative ions which pervades the whole channel. Ash and fume particles pick up ions and become negatively charged as soon as they enter the field; electrostatic forces cause them to drift to the collecting walls to build a deposit which is removed through a hopper by periodic rapping. This sounds beautifully simple but, in practice, hardened deposits that resist vigorous knocks can build up and eventually disturb the field and impede the flow of the corona current.

As with many electrostatic devices the efficiency rises sharply with voltage but, since the electrodes cannot be kept clean, flashover depends on many uncontrollable factors. It is thus necessary to adjust the voltage empirically to the maximum without sparks which would disturb operation for as long as about one second. Today the operator at the control panel is replaced by an electronic regulator which counts the sparks and holds the voltage at the brink of flashover. The value of such an automatic control has been demonstrated by practical measurements.

Electrostatic precipitation is also important in the steel, cement, and chemical industries, where flue gas outputs are also very large. It is little used for smaller gas flows, except in the case of room air conditioning in which special designs are necessary to minimize ozone formation.

Electrostatic spraying, dusting, and powdering, which have aroused great interest since 1950, are similar in principle to precipitation but require very different equipment. The paint, varnish, plastic powder, or pesticide is electrified by flowing past an electrode at a d.c. voltage of up to 150 000 V. The electrode shape and the charging mechanism differ according to the material handled. Liquids can be atomized and

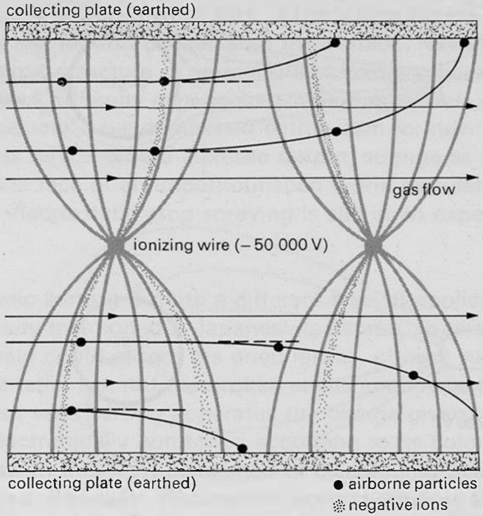


Figure 5
Electrostatic precipitation of particles in a gas flow can be achieved by spraying them with negative ions. The charged particles are then attracted by earthed collecting plates. The electric field is present everywhere but the ions are emitted in narrow beams from the sharp edges of microscopic defects on the wires.

charged simultaneously but solids must be powdered beforehand. For liquids the spinning disk or cylinder, which acts as both centrifugal atomizer and charging electrode, is being replaced in many applications by a low pressure air gun with a charging nozzle. The highly charged droplets repel each other but are attracted by the nearest earthed object upon which they build a uniform layer. Paint waste may be as little as ten per cent, even for difficult shapes like grids and bicycle frames. The manpower requirements of the process are insignificant, and it lends itself to complete automation.

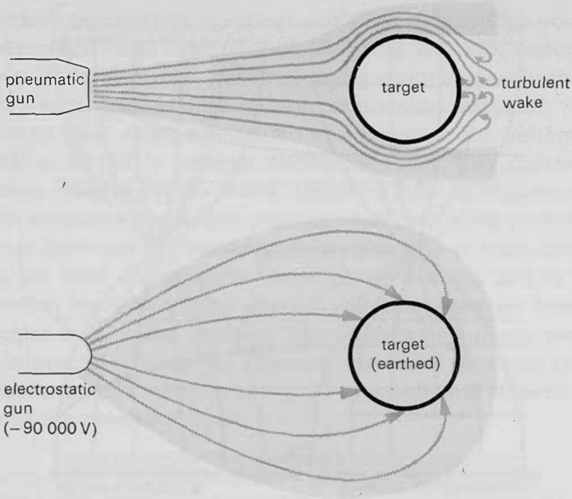


Figure 6
Electrostatic spraying causes charged droplets to travel along paths which lead to every part of the target. In contrast, the pneumatic sprayer blows most of the lighter droplets past the target.

Plastic powders can be sprayed in a similar way, and adhere to a metal surface for minutes and even hours provided they are good enough insulators to keep their charge. Subsequent heating in an infrared oven turns the powder into a very hard, protective layer. Though more expensive than painting, the process is very promising when durability is at a premium.

There are innumerable other applications of this nature, for every spray or powder can be usefully electrified – talcum powder in the rubber and cable industry, water sprays in paper mills to maintain correct moisture content, aluminium powder on steel plates to prevent oxidation during hot rolling, and many others. In the past 20 years several attempts have also been made to use charged pesticides for crop protection. Results have been mainly satisfactory: the particles strike the leaves at much higher velocity, improving adhesion even after loss of the charge, and they are more evenly distributed

– particularly on the reverse side of the leaves where most pests find shelter and usually start their attack. Nevertheless the economic structure of agriculture in Western Europe, and the perpetual difficulty experienced in finding reliable customers for its products, have dampened enthusiasm for new techniques which would increase output, strange as this may seem in the face of an undernourished world population. Certainly electrostatic crop spraying is still in an experimental stage.

Electrostatic sorting is quite a different type of application. I have already mentioned a Japanese tea sorter. In this case the electric field could attract the dried leaves, of high surface-to-weight ratio, but not the broken stems and wood fragments. In a British seed sorting apparatus the charge given to each seed is electronically controlled according to its colour and brightness. In this way, blemished or discoloured peas can be removed. Basically, electrostatic sorting requires that the forces acting on the individual components of the mixture should vary according to a convenient physical criterion, such as surface-to-weight ratio in the first case or colour in the second. Again there is a potentially unlimited range of applications; but much research and engineering remains to be done to take full advantage of this potential.

At present electrostatic sorting is used mainly in mining. Longer established flotation and magnetic techniques suffer from shortcomings: wet methods are inconvenient in frigid or arid climates and magnetic sorting is limited in application. These and other reasons have directed attention to electrostatics, and considerable development has taken place in the past 20 years.

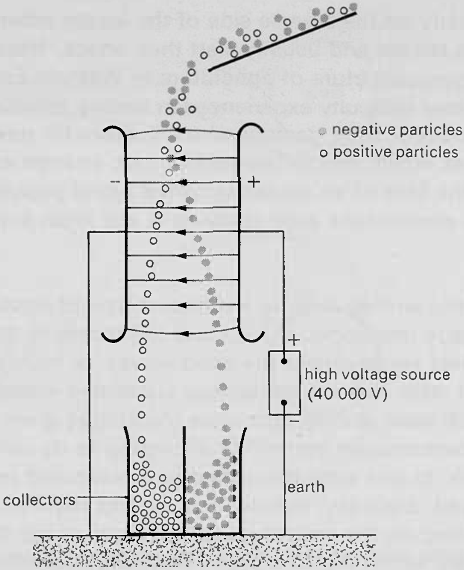


Figure 7
Electrostatic sorting depends upon differences in factors such as contact electrification or surface area between the mixed particles. These differences cause the particles to follow dissimilar trajectories when falling through an electric field (indicated by arrows).

A typical ore is made up of two minerals, A and B, liable to electrification by mutual friction. After grinding to free the crystals the ore is made to fall through a horizontal electric field. Because mutual friction produces charges of opposite sign, the A and B crystals are deflected in opposite directions away from the vertical and are thus separated. This procedure still works with more than two components if the valuable material takes a sign opposite to the others, as in a calcium phosphate/limestone/silica mixture. Friction causes both the limestone and silica to become negative but the phosphate positive.

When friction electrification is of no avail electrostatic separation can still be achieved by resorting to the wide differences in conductivity exhibited by most minerals. The ore is fed on to a rotating earthed cylinder while a wire at 30 000 V d.c. sprays it with ions. The charged particles adhere to the cylinder by electrostatic attraction but, as soon as they leave the ionized region, they begin to lose their charge according to their conductivity. Good conductors neutralize after a fraction of a second and fall away, but insulators such as silica remained attached until they are removed by a brush.

Accelerators

Electrostatic accelerators play a considerable role in nuclear physics – both fundamental and applied – and also did much to revive interest in electrostatics. The essential part of such an accelerator is an evacuated insulating vessel, usually referred to as the 'tube', where ions (mostly positive) encounter a negative potential gradient and thus reach very high speeds before impinging on a target where they trigger nuclear reactions. A common ion source is a self sustaining discharge in rarefied hydrogen or deuterium confined in a separate bulb communicating with the main tube by a pinhole. An auxiliary voltage shoots ions through the pinhole while neutral molecules having only thermal velocities escape more slowly so that the vacuum can be maintained by powerful pumps.

Principal data for such a device are the accelerating voltage and the ion beam current, which respectively determine the ultimate kinetic energy of the ions and the number of ions per second. Many electrostatic accelerators yield beam currents in excess of one milliampere, or more than 10^{16} ions per second. The accelerating voltage has an essential bearing on the likelihood of nuclear reactions and, therefore, on the final yield of neutrons or other reaction products for a given beam intensity. Most accelerators require voltages in excess of 100 000 V and the latest Van de Graaff machines go up to 10 million volts. Such voltages cannot be safely insulated except in tanks pressurized at 700 to 2100 kN m^{-2} and this makes maintenance complicated. Accelerators working below 1 000 000 V can usually operate in free atmosphere.

Neutron generation is the most important application of small accelerators ranging from 100 000 to 600 000 V. Provided the target is properly charged with tritium (a radioactive hydrogen isotope) the neutron yield from a deuteron beam 4 mA in intensity and accelerated under 400 000 V can reach 10^{12} neutrons per second. Such neutron sources are very useful for nuclear research, reactor engineering, and activation analysis. Portable units rated at 2 mA and 150 000 V can give 10^{11} neutrons per second, sufficient for several applications such as quick activation analysis of steel samples and training in nuclear technology.

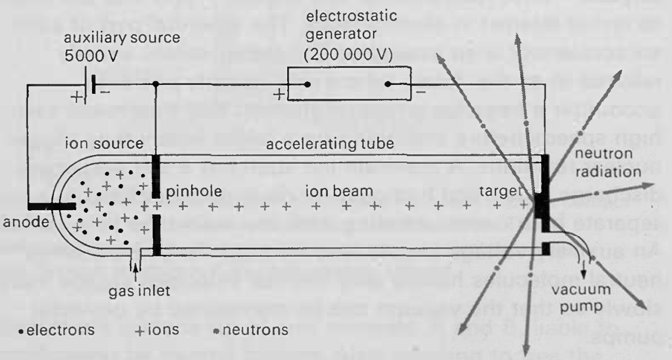


Figure 8
An electrostatic accelerator is illustrated as a neutron generator. Deuterium gas is admitted to the ion source and ionized by an auxiliary discharge, the electrons being collected by the anode and the positive deuterons being expelled as a beam through the pinhole. The deuterons impinge on the target, which is seeded with tritium, and a few of them (about 1 in 10^5), making head-on collisions with tritium nuclei, knock fast neutrons out of them.

Electrostatic accelerators in the 1–10 million volt range are mostly of the Van de Graaff type in which an accelerating tube and a high voltage electrostatic generator are combined into a single compact unit insulated by gas pressure in a steel tank. No external connection to a high voltage source is required and insulation problems are easily solved in an elegant way. Van de Graaff machines have generated the highest man-made voltage throughout the past 30 years.

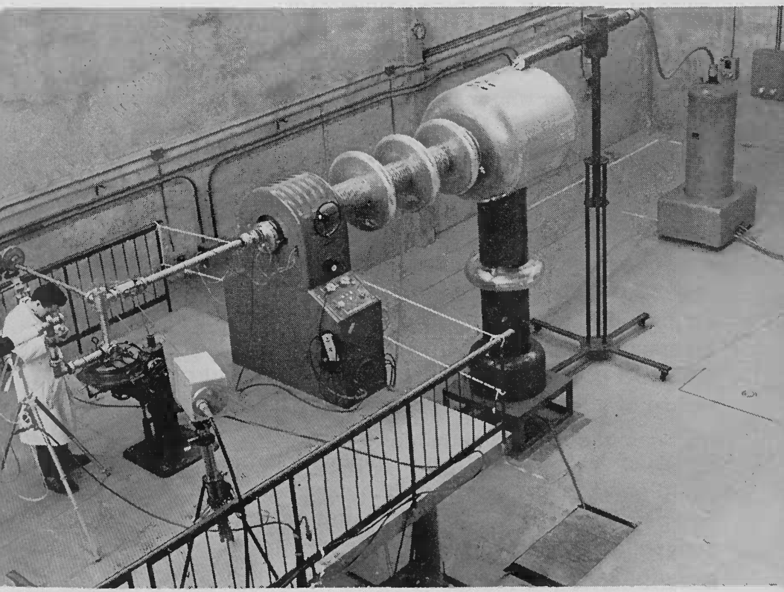


Figure 9
The horizontal ion accelerator is served by a 600 000 V electrostatic generator contained in the vertical tank on the right. Supply of 4×10^{-3} ampere is fed through a high voltage cable and damping resistor. Ions are accelerated in the vacuum tube with three toroidal intermediate electrodes at cascaded potentials given by a coiled resistor. The slender tube on the left is a metallic extension which is used to carry the target material.

Electrostatic engineering

It might seem obvious that a field with such wide applications as electrostatics would be included in all basic engineering curricula, as are thermodynamics or electromagnetics. In fact virtually no significant training in applied electrostatics has been attempted in basic engineering education. This neglect has given electrostatic engineering some unusual features. Many advances have sprung not from organized research but from an old fashioned 'spirit of invention' or sheer amateurism.

Nevertheless, electrostatic engineering has already had a considerable economic impact. In many applications equipment has often paid for itself – by savings in raw materials or manpower – in less than a year and sometimes in a few weeks. Paradoxically, the electrostatic equipment industry seldom makes easy money. Almost every machine it sells replaces a much greater quantity of conventional gear as very few electrostatic appliances are needed to handle enormous industrial outputs. At present, a conservative estimate of the savings realized by electrostatics in Western European industry is £10 million a year, which is almost the total cost of the equipment involved. Such returns are typical of a really new physical agency being put to work.

However, all but an almost negligible fraction of industry has yet to become 'electrostatically conscious'. Most of its powders, sprays, pesticides, and ores are still handled by conventional mechanical means, with consequent wastage, pollution, and health hazards. The rank-and-file engineer is invariably unable to appreciate the value of electrostatics because electrostatic forces have not been properly explained to him. This neglect has undoubtedly retarded the industrial acceptance of electrostatic equipment.

This situation has markedly improved in the past few years and I anticipate a considerable expansion and diversification of applied electrostatics in the near future. The financial savings and the improvement in living standards resulting from decreased environmental pollution to be expected from this expansion are truly enormous, and warrant substantial backing by government and economic agencies. In the long run I can even visualize the advent of powerful electrostatic energy converters, capable of displacing electromagnetic machines in certain applications. The theoretical arguments which have been produced in favour of such a development have doubtless cleared the way to some extent, but practical engineering attempts to produce high power machines have so far been unsuccessful.

This means that the problem must still be approached from the physical side. Potentialities for stronger and more reliable electric fields are known to exist in vacuums and in liquids. Purified nitrobenzene, for instance, has been shown to withstand a field strength of 400 000 V cm^{-1} with a dielectric constant of 36; this corresponds to an electrostatic pull of 250 kN m^{-2} which is energetically equivalent to a magnetic field of 0.8 tesla, a fairly good figure for the gap of a conventional machine. Such a field strength in an engineered structure would allow electrostatic converters to reach much higher specific powers without resorting to complicated arrangements of interleaved sectors.

Such converters would exhibit quite unusual features and could at the same time solve numerous problems which have plagued electrical engineers for nearly a century. For example electrostatic motors could run directly on high voltage, produce a great deal of reactive power, and yet be switched on and off irrespective of synchronism, speed, or excitation, without giving rise to any surge of current or power. Chokes and rheostats would not be needed and electrical engineering would achieve a simplicity and flexibility that it has so far conspicuously lacked.

Today electrostatic engineering can no longer be regarded as unscientific gadgeteering outside the stream of modern technology, but must be incorporated as a regular section of electrical engineering. Its scope is truly enormous, since it is not restricted – as is electromagnetics – to a specific kind of material. Electrostatic forces are supreme as far as divided matter is concerned, and this is of great value to most segments of industry and agriculture. In the future they may even outstrip electromagnetics in strength as well as in ubiquity and flexibility. After all such an upheaval in technology would be perfectly consistent with the known precedence of the order of forces in nature.

Professor Noel Joseph Felici is Professor of Electrostatics at the University of Grenoble, France. His current research embraces electrostatic generators, discharges in compressed gases, and superconductors, mainly supported by the Centre National de la Récherche Scientifique. He is a consultant to the Commissariat à l'Energie Atomique and his team co-operates with the *S. A. de Machines Electrostatiques*.

Acknowledgements

Figures 3, 4, and 9: Tunzini Sames, Grenoble, France.

Wilem W. Frischmann

Tall buildings

Increasing the height of buildings saves space on the ground and may bring other benefits. Today it is technically possible to build a city for half a million people on a foundation 150 metres square.

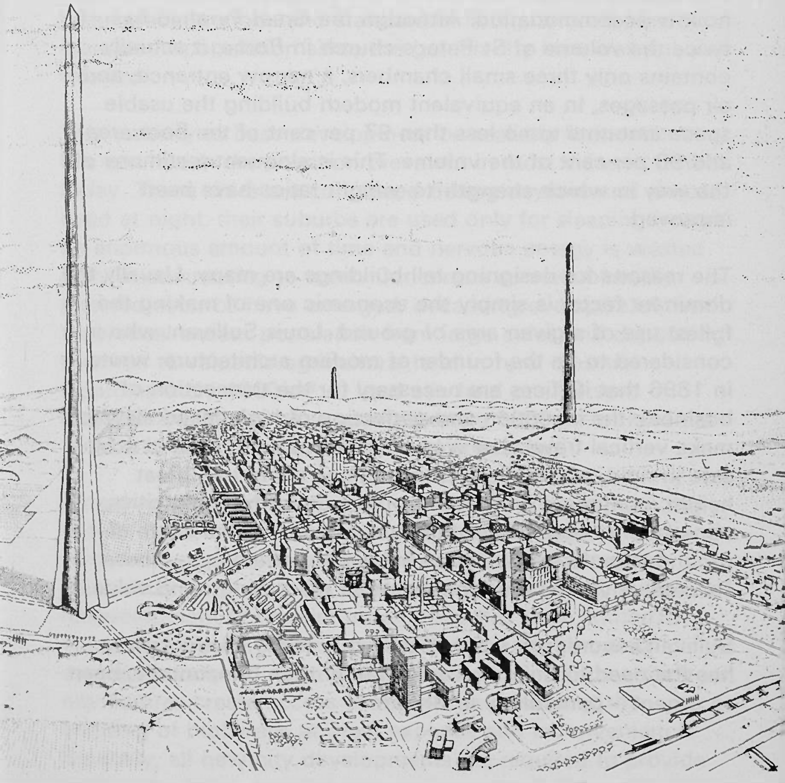


Figure 10

The earliest tall buildings were essentially solid mounds of a pyramidal form. Such structures reached their zenith with the impressive Egyptian pyramids dating from about 2700 BC onwards. Largest of all was the tomb of King Khufu – the Great Pyramid at Gizeh. Originally this reached a height of some 145 m, equivalent to a modern building of 42 storeys. To build it slaves hewed and placed in position some 2 300 000 blocks of stone each weighing between $2\frac{1}{2}$ and 15 tonnes.

But this method of construction hardly lent itself to usable hollow accommodation. Although the Great Pyramid has twice the volume of St Peter's church in Rome, it actually contains only three small chambers, a narrow entrance, and air passages. In an equivalent modern building the usable space amounts to no less than 97 per cent of the floor area and 90 per cent of the volume. This is eloquent testimony of the way in which strength-to-weight ratios have been improved.

The reasons for designing tall buildings are many. Usually the dominant factor is simply the economic one of making the fullest use of a given area of ground. Louis Sullivan, who is considered to be the founder of modern architecture, wrote in 1896 that 'Offices are necessary for the transaction of business: the invention and perfection of high speed elevators make vertical travel that was once tedious and painful now safe in rigid, economical constructions rising to a great height; continued growth of population in the great cities, consequent congestion of centres and rise in the value of ground stimulate an increase in number of storeys – these successively piled one upon another react on ground value.'

Sullivan's words are emphatically true today. Never before has the need to build vertically, rather than horizontally, been so urgent – particularly in Britain.

Town and city planning

Plans for the development of British towns and cities since the 1947 Town and Country Planning Act have unfortunately been closely based on the humanitarian movement associated with 'garden cities' developed since World War I at places such as Letchworth and Welwyn. The Act was prepared apparently without reference to the part that tall building and the multiple use of land can play, although these have been features of towns in other countries for many decades. The policies laid down do not take into account advances in science and technology, the enormous changes that have taken place in mode and standard of living, and the vast increase in population since the garden city idea was put forward.

This inability to base development on realistic foundations has resulted in the situation prevalent in most cities in Britain today. These cities are congested during daytime and almost dead at night; their suburbs are used only for sleeping in and an enormous amount of time and nervous energy is wasted daily in commuting to work. Cancerous sprawls of badly planned suburbs have emerged, decanting overspills of the theoretical excess population from large cities and depositing them on to valuable agricultural land in rural country. Approximately 4000 m^2 of land can feed one person. England and Wales have approximately 3000 m^2 of land per head and barely half of this is food producing.

Moreover, the introduction of 'green belt' areas to limit this outward spread has created an artificial shortage of land and forced the population to develop land beyond the green belts, thus increasing the necessity for commuting. Density of redevelopment of derelict urban areas has been limited to approximately 20 persons per 1000 m^2, and the urban average is only 10. Overspill dormitory towns kilometres from city centres create access problems, necessitating the building of highways at great expense and inconvenience. Similarly, all new city developments are required to provide parking facilities, encouraging car usage and filling the city not only with people but also with vehicles.

If the garden city concept were universally accepted one fifth of the area of the United Kingdom would be built on. But an attempt is at last being made to ease the problem by increasing the normally accepted plot ratio – total floor area divided by site area – by building towers of flats or offices up to 135 m in height. What is considered the most advanced British town plan is that of the Barbican in the City of London, where an attempt is being made to combine both types of accommodation in one development so that people can live within walking or elevator distance of their work. However the majority of the working population are unlikely to be able to afford to live there.

Looking down from an aircraft as it comes in to land, there is no escaping the conclusion that advanced architecture has barely touched our lives. The ugly earth-hugging sprawl of each town and city, the brick-box-cave architecture, conform to a pattern handed down from the Babylonians.

The only solution appears to be an intensive use of land – not the construction of tall blocks of flats which are only an extension of bungalows built one on top of another with all the inherent social disadvantages. A person, especially an old age pensioner, living on say the 18th floor of a block of flats sees on average four different people each day across the landing and in the lift. For social contact this person might be involved in a lift journey (if the lift has not broken down) and a long walk through traffic-congested streets impossible to cross, perhaps down to the shopping centre. One can easily understand the sense of isolation, frustration, and desperation of living in tall blocks of flats.

The pattern could change – three-dimensional cities built in the sky filled with human activity of every kind providing a balanced community. The present commuter population could then be rehoused within existing town boundaries, whilst derelict urban areas were reclaimed.

As we shall try to demonstrate, this is now technologically feasible and it is socially desirable that such cities should be seriously considered.

Building high

Prior to about 1870 very few buildings contained more than five or six storeys. To go higher was structurally difficult and the number of stairs to be climbed became excessive. But at that period the passenger lift became available and nine or ten storeys became an acceptable limit. A typical tall building of this era is a block of flats designed by Norman Shaw and built in Kensington, London, in 1879. Chicago was rebuilt following a disastrous fire in 1871, and here arose the ten-storey Home Insurance Building in 1883 and the 16-storey north half of the Monadnock Building, 60 m high, in 1891.

To comply with local by-laws these structures were built with solid masonry walls which were both load-bearing and fire-resistant. This adherence to established methods of construction restricted further increases in building height, but the emergence of metal structures such as the 300 m Eiffel Tower was a pointer to the future. At the turn of the century authorities accepted the revolutionary idea of designing a tall building around a metal skeleton bearing all the main loads, and relegating the exterior walls to the function of mere non-structural cladding. Once this was done the 'skyscraper' became possible.

On Manhattan Island in the city of New York a combination of high ground values, tremendous demand for office accommodation, and a rock foundation led to dramatic increases in building height. The famous skyline began to take shape with the 20-storey Battery Place Building, the 38-storey Metropolitan Life Building of 1908, and the 60-storey Woolworth Building of 1913 – a jump to 241 m.

After World War I a remarkable spirit of competition was created in the United States, each structure seeking to outgrow its nearest competitor as tall buildings became the symbol of American affluence. Edifices such as the 71-storey Manhattan Building and the 77-storey Chrysler arose in New York, while Chicago and Philadelphia also entered the race.

The greatest achievement of this thrusting skyscraper epoch was the Empire State Building completed in New York in 1932. The basic 102-storey tower rose to 381 m and the post-war addition of a television aerial mast has increased the height to 448 m. The building incorporates shops, offices, restaurants, a bank, Turkish baths, and clubs, and contains 65 lifts. Another notable New York development of this era is the Rockefeller Center. This is designed on a concept of 'a city within a city' and consists of the 70-storey RCA building, the 36-storey Time-Life Building, the 41-storey International Building, and others, grouped around a piazza and linked at several underground levels. They contain a large theatre and cinema, stores, offices, and radio and television stations used by approximately 150 000 people daily. All these buildings have steel skeletons built up on a steel cage system and protected against fire by encasement with concrete, bricks, or stone, using reinforced concrete in foundation slabs and stairs where extra fire resistance is particularly essential.

Since World War II many towers, including the 30-storey United Nations Secretariat, the 24-storey Lever House, and the 59-storey Pan-Am Building have retained the steel skeleton form of construction, but with new cladding materials: stainless steel, bronze, copper, zinc alloys, glass, plastics, various ceramics, thin skins, and non-load-bearing precast concrete mullions.

In Europe – particularly in Britain – and now also in the United States there has been a dramatic development in cement manufacture and concrete technology, coupled with advanced constructional techniques and on-site mechanization by cranes, hoists, and pumps. As a result tall buildings have now been designed and constructed not in steel but in reinforced concrete. Examples include the

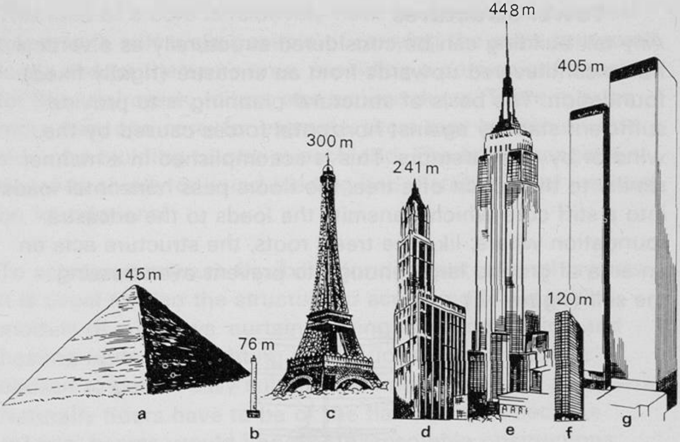


Figure 11
The evolution of tall buildings is portrayed in these sketches drawn to a common scale. The bulk of the Great Pyramid (a) is impressive but this 2700 BC structure is virtually solid. The Torre Asinelli (b) at Bologna typifies many masonry-walled towers of the 12th century. Although hardly an inhabited structure the Eiffel Tower (c) of 1887–9 demonstrated the dramatic increase in height which could be achieved with a wrought iron skeleton. The Woolworth Building (d) of 1913 has a steel frame faced in stone, as does the Empire State Building (e) of 1932. Contrasting post-war designs are Centre Point, London (f), the highest precast concrete structure, and the projected all steel World Trade Center (g).

Pirelli Building in Milan, the Millbank Tower and Centre Point in London (all approximately 120 m in height), the 179 m Marina City in Chicago, the 50-storey American Hotel in New York City, and the 192 m tower at Place Victoria, Montreal.

Even larger, the projected New York World Trade Center would rise 110 storeys to 405 m. Its design specifies twin towers each with a central steel core and external load-bearing steel mullions. But in my opinion a sad feature of this building is that the designers are using a plan type developed for, and most suitable for, reinforced concrete construction.

Tower structures

Any tall building can be considered structurally as a vertical frame cantilevered upwards from an encastré (rigidly fixed) foundation. The basis of structural planning is to provide sufficient stability against horizontal forces caused by the wind or by earth tremors. This is accomplished in a manner similar to the action of a tree: the floors pass horizontal loads into a stiff core which transmits the loads to the encastré foundation where, like the tree's roots, the structure acts on an area of ground large enough to prevent overstressing the soil (*figure 12*).

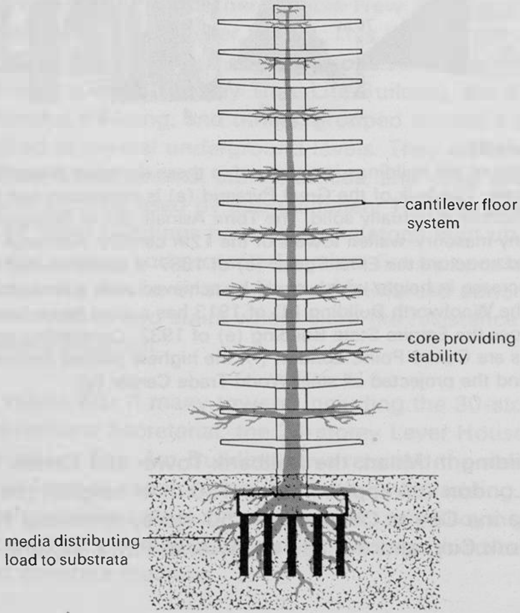


Figure 12

Structural similarity between tall trees and tall buildings of the core type.

This idea of a core is relatively new. It consists of a vertical tower with stiff walls capable of carrying the main structural loads and, at the same time, providing a convenient housing for lifts, staircases, toilets, and service ducts. There may be more than one core: for example, there may be two on one side of a building or one at each end. The best arrangement depends chiefly on wind loading and distribution of pressure on foundations.

To achieve maximum flexibility for different internal layouts it is usual to plan the structure to accommodate a basic module of windows, curtain walling, air conditioning and heating units, and lighting installations. This module is chosen to permit easy subdivision into convenient room sizes. Naturally floors have to be of the flat slab type because internal beams would impose unacceptable obstructions. At the same time there is great pressure on the architect to incorporate floors of uniform minimum depth, and a span-to-depth ratio of 30 to 40 is today typical.

Since the floor element stays basically constant from top to bottom it is amenable to mass production either as a hollow, ribbed steel deck or as precast concrete panels. Another arrangement is to use deep lattice trusses spanning 18 or 21 m to carry cellular metal decking about 5 cm deep topped by lightweight concrete and with an underside (ceiling) sprayed with vermiculite for fire protection. In the latter case concrete at 2400 kg m^{-3} has now largely given way to foamed slag, expanded clay, or pelletized and sintered pulverized fuel ash, all at around 1600 kg m^{-3}.

In residential blocks the smaller cross section of core required to house the lifts and services may call for additional stability provided by stiffened concrete shear walls. In office building it is often found that the optimum design of structure uses load-bearing perimeter columns, and in this case the precast concrete mullion is very suitable. These may be precast in sections three or four storeys high or cast as storey height H frames incorporating lightweight concrete spandrel walls.

At an early stage in the design the engineer must begin calculating the effects of wind, vibration, and, in some countries, earthquakes. Air tends to flow down on the windward side of a tall building, often causing strong eddies at pavement level to the annoyance of pedestrians, and upwards on the leeward face to cause problems of rain and snow penetration through the cladding. Large elements of cladding in recent buildings have frequency and damping much lower than previously, and without careful design are prone to gust-induced sway.

Aerodynamic instability of structures in steady winds arises from 'galloping' and vortex shedding. Galloping instability, encountered on iced transmission lines, is unlikely in buildings except in cases of very low damping or winds of hurricane force. But the second type of instability is of great significance to the design of tall buildings, as a vortex-shedding frequency at normal wind speeds can coincide with the natural frequency of the tower. If this is the case it allows resonant vibration.

Vibration has to be considered not only in relation to the stresses induced, and the resulting fatigue life of the structure, but also from the viewpoint of the people in the building. Human beings are very sensitive to vibration, and at the frequencies encountered the maximum amplitude they can comfortably tolerate is considerably less than the limit imposed by building stress considerations. Numerous workers have published limiting values for accepted interior vibration and acceleration. What is important for the designer of tall blocks is that the accepted amplitude of transverse vibrations must be considerably less than that anticipated from static wind loading.

Earthquake forces result from movement of the ground both vertically and horizontally, but the vertical component usually has a negligible effect on structures with such an excess of vertical strength. Earthquake design forces are at present empirical, based primarily on the demonstrated performance of structures in earthquake zones. Some countries require that, in addition to the normal wind forces, structures shall withstand a lateral seismic force along either main axis equivalent to from 1 to 13 per cent of the weight of the building. Walls and partitions must be anchored to resist a force of 20 per cent of their weight and members such as parapets, ornamentation, and cladding are usually required to be anchored for 100 per cent of their weight.

Structural behaviour is investigated by advanced theoretical analysis. Much of this work is now simplified by the use of digital and analogue computers; standard programmes are available for the design of frames, grids, piling layouts, and similar basic variables. For a complex structure it is also advantageous to carry out analysis using models varying from 1/5th to 1/50th scale constructed from Perspex and other plastics, cement mortar, steel, aluminium, and other substances. The expense of such tests is more than offset by the saving in structural costs from using a more refined factor of safety.

Foundations

Foundations are of obvious importance, especially to a tall structure. In the past foundation difficulties often imposed a major restriction on the buildings erected. With greater knowledge and experience of soil mechanics, large buildings can now be founded economically on relatively poor soils. Laboratory tests on samples of soil make it easier to determine the probable distribution of stress, the settlement, and the optimum substructure.

In proportioning foundations, the allowable bearing pressure on the ground must be chosen to provide an adequate factor of safety against shear failure, and to ensure that both overall and relative settlements are tolerable. On non-cohesive soils, such as sand and gravels, water can move freely and so the settlements are largely completed by the end of construction. On cohesive soils, such as clays, the spaces between the particles are small and the movement of water under pressure can take place only very slowly. Theoretically, consolidation can continue almost indefinitely, although it is normally imperceptible after eight or ten years. This long, continued settlement causes much trouble with structures and finishes.

In big cities considerable experience has been acquired in founding tall blocks, in forming deep basements, and in stabilizing neighbouring buildings against possible lateral and vertical movements. The London area is underlain by a great thickness of clay which increases in compressive strength with depth. *In this city the choice of foundation for* tall blocks lies mainly between a raft, a buoyant foundation, and a pile foundation, or a combination of these.

The function of a raft is to spread the load over a large area in order to reduce the bearing pressure to the allowable limit. Normally it is made stiff enough, either by itself or in combination with the substructure, to distribute the load and make differential settlements acceptable.

The principle of the buoyant foundation is that of excavating earth weighing approximately as much as the building above in order to reduce the average settlement to a minimum. In practice the difficulties of forming a sufficiently deep excavation make a fully buoyant foundation expensive. Usually a compromise has to be accepted; this can take the form of a raft at, say, the sub-basement level extending well beyond the plan outline of the tower block, or a raft placed at a shallower depth used in conjunction with piles which transfer the loads into stiffer and less compressible soil deeper down. Large diameter augered piles, transferring loads partly in friction and partly in end bearing, are very economic; their bases can be enlarged by under-reaming.

Services

Essential services need most careful consideration. A normal rule of thumb for lifts is one person to each 10 m^2 of floor, of whom one in six (one in 20 for flats) has to be moved in an elapsed time of five minutes. Waiting time should not exceed 30 seconds. Very large buildings are now divided into zones each served by their own express lifts and containing slower lifts for 'local traffic'. Water systems are complicated by the need to avoid excessive pressure behind draw-off points while supplementing mains pressure with booster pumps at upper levels. Drainage involves independent discharge from lower levels to prevent back-flooding in the event of stoppage at the bottom of the pipe. But for refuse disposal, the possibilities include chutes to containers at ground level, sink grinders from which the refuse is piped by water, incinerators, and dustbin collection.

Upper levels of tall buildings are exposed to sun, wind, and rain. Average wind speed is substantially higher than at ground level and natural ventilation is normally unsatisfactory. Moreover, internal heat gains, mainly from lighting and occupants, and solar gains from direct and indirect sunshine, are such that simple mechanical ventilation cannot always provide comfort. Air conditioning is necessary, ventilating with filtered air, cooling, heating, and either humidifying or drying as required. Moreover ducts supplying air at a speed of 15–20 $m\ s^{-1}$ require high fan pressures and so fan silencers are essential.

Industrialized methods

Industrialized methods of building have now become almost universal for large structures. At the beginning of the decade, a start was made by precasting some concrete elements off the site. Since site-cast concrete is cheaper, off-site precasting had to show multiple advantages. It reduces construction time and site labour – especially of skilled craftsmen – and guarantees a finish, accuracy, and uniformity that cannot be achieved on site.

In general the infill cladding which forms the main exterior surface should be capable of easy erection, cleaning, and maintenance without scaffolding. A difficulty with concrete cladding is anchoring it to the main structure. Problems with tolerances often arise because of cumulative inaccuracies in the concrete structure and the precast cladding units. Nevertheless precasting can now be accomplished by mass production conveyor-line techniques incorporating various finishes, and ducts for plumbing, electrical, and other services. I see no reason why units comprising complete rooms and services should not be preconstructed in a factory under quality-controlled conditions and then brought to the site. These modules would then require only connection of services.

In Britain several 'hanging' buildings are in the course of design or erection. This form of construction provides a light and elegant structure with a column-free open space at ground floor level. A reinforced concrete central core is first constructed in sliding shuttering – a 60 m core can be completed in four weeks. Cantilever arms, either in steel or prestressed concrete, project from the top of the core and ties are hung from them. The ties can be steel cables, flat bars, or prestressed concrete. The floors are all cast on top of each other at foundation level while the core is being built, hoisted to their final position, and connected to the core at

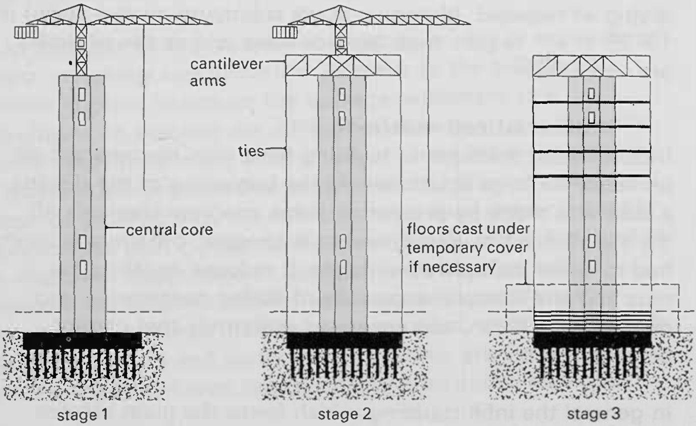


Figure 13
The erection sequence of a typical core-type building starts with construction of the core itself surmounted by a tower crane. At the top are added cantilevers from which ties are hung to support the floors cast at the base and then raised.

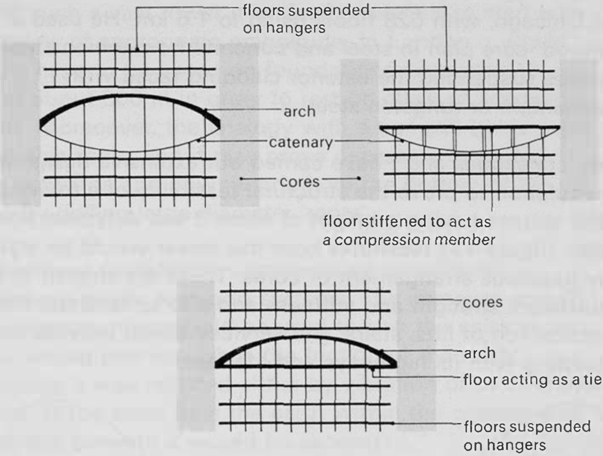


Figure 14
Floor hangers may be suspended from the top of a building core as illustrated in figure 13. Alternatively, support may be provided at intervals by additional cantilevers or, in the case of a building with two cores, by bridge structures consisting of arches, catenaries, or stiffened floor members as shown in these three side-elevation sketches.

the centre and to the hangers at the perimeter. The finishes can be applied and even partitions erected at ground level. Alternatively the floors can be cast, from the top downwards, on moveable shuttering hung from the structural core. Some economic advantage may be gained by inserting cantilever arms every 20 to 30 floors, to reduce the size of the ties and their extension under load.

Three-dimensional cities

This basic principle of construction can be used to erect buildings of unprecedented height, rising some 3 km from the ground as I indicated earlier. Such vertical development opens the way to the complete city housed in a single structure.

Frank Lloyd Wright, the great genius of American architects, was one of the first to propose a very tall building. In 1956 he published his concept of an office building, for erection in Chicago, with 528 floors rising to 1.6 km. He used a 'tripod' core plan in steel and concrete from which steel cables suspended the exterior cladding faced with aluminium or tungsten steel.

My colleagues and I have carried out extensive design and calculation to prove the structural feasibility of a tower of 850 storeys having a height of some 3 km. A typical floor plan (figure 15) illustrates how this tower would be stabilized by judicious arrangement of cores. These are shaped to have maximum strength and stiffness and also to facilitate the vertical run of lifts, stairs, and services which provide the building with its highways and lifelines.

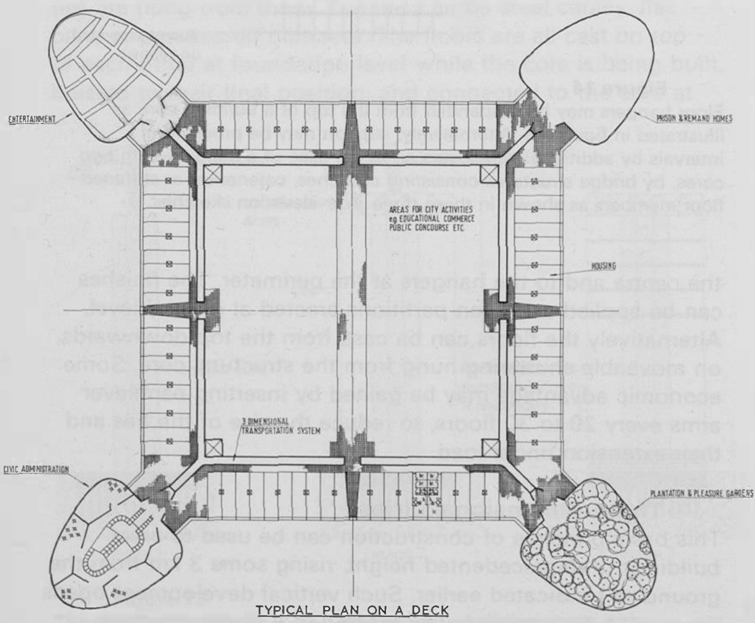


Figure 15
A typical plan of a deck of a tower city.

To build such a vast tower it is clearly essential to start with foundations of appropriate magnitude. In London it would probably have to be based on foundations excavated to a depth of about 300 m in order to rest on chalk of adequate strength. Moreoever, the analogy with a tree still holds good even on this scale, and it may prove necessary to go even further outwards and downwards from the bottom of the caisson by drilling large diameter 'roots'.

This enormous foundation caisson could incorporate underground public transport systems. The immense weight of the superstructure would be used to sink the caisson into its final resting position while the friction of the earth surrounding it was reduced either by vibration or by chemical injection; at the same time the earth within the caisson and immediately beneath it would be excavated.

From the foundations cores of reinforced concrete would rise vertically to provide the basic skeleton. By careful control of the aggregates and manufacturing processes, it would be possible to maintain a compressive strength of about 10^8 N m^{-2}. Cost of building in such dense and strong concrete would fall considerably if the technique became widely accepted.

Every 30 floors a balanced system of arches and catenaries would be inserted at service floors. The 15 floors immediately below would be suspended on hangers of high tensile steel while the 15 floors above would be supported on columns of similar material. The floors would be lattice trusses 1.2 m deep to accommodate horizontal services and would need little support by intermediate columns. The cladding would be precast concrete infill panels made up into 'brises solaire' to reduce direct glare through the inserted glass panels. It would be set back 1.2 m to give the occupants a greater sense of security. Air conditioning would maintain pressure near sea level value.

This building would be large enough to form a vertical 'super city' in the 'total building' sense. It would provide accommodation for living, working, education, shopping, and every other amenity for community life for some 500 000 people. The ground space would be left free for gardens, lakes, recreation grounds, the production of food, and other activities.

Cities will be built round a concrete centre core 180 m square and have up to 850 levels, each with housing units on the outside face. Each unit will have its own garden, and a breathtaking view of the countryside far below.

The 180 m square area within the central core will be given over to a different activity at every level. Some will be factory sites; others will be office complexes. Many will be shopping precincts, or 'silent' levels with schools, hospitals, churches, and libraries.

At regular intervals up the tower, squares will be recreational areas – piazzas, perhaps, with fountains, cafés, and bars, or spacious indoor parks.

No one living within the city will need to travel for more than three minutes to get to work. Everyone will have an independent share in a public transport system, able to programme his or her own movement to any part of the structure or out of it.

One hundred families will live on each level. They will come together naturally as a community, served by shops supplying immediate needs – a dairy, a newsagent, a launderette.

Every twenty levels will be grouped together, with their own supermarkets, general stores, community and health centres, churches and chapels, police stations, libraries, and primary schools. These will be 'villages' within the tower with 6000–7000 people governing their own local affairs.

Ten such village groups will make a city district. Each district will have its own cinemas, hotels, offices, newspapers, civic buildings, secondary schools, technical colleges, swimming pools, sports centres, and industry.

Home life will have little in common with that in today's tower blocks of flats. Remember, that to walk from one side of a tower city to the other will be like walking across Trafalgar Square. The person in the tower city will make friends among at least 100 families. Physical movement will not be restricted to four or five flats and the landings between them. In a tower city there will be opportunities for far greater local travel, all of it under cover.

In housing areas, inhabitants will, as now, get exactly what they pay for. These areas will begin as rudimentary shells, glass-walled on the outside. Homes will be planned into them just as houses are planned into a new stretch of road – executive homes, bed-sitters, three-room flats, maisonettes.

Between the living areas and the glass wall will be greenhouses. The higher up the tower, the greater the variety of plants that will thrive in the strong sunlight being filtered in.

Each home will be connected to the tower's central power and water supply, waste disposal plant, and air conditioning/pressurization system. At all living levels temperature, humidity, and pressure will remain constant. Clothing conventions will take a hard knock. Summer and winter, spring and autumn will be alike in the tower.

Movement between levels will be by personal car-sized capsules, clipped on to a central power line and programmed by their occupants. A citizen living on level 467 may work in an office on level 773. His capsule, working on the linear-lift principle already known to engineers, will travel vertically and horizontally, delivering him to any destination in not more than three minutes.

People and goods will move across each level by motorized footways. Heavy goods will be delivered by vertical lifts operating within the central core.

Huge underground car parks outside the city will accommodate the city's 150 000 private automobiles: motorists will be able to join any one of the motorways that will pass near the tower within minutes of leaving their homes.

The future

Everything connected with the three-dimensional city concept is exciting and challenging. For the first time, city dwellers' domestic, social, and working habits are going to be re-thought. Conventional city problems of congestion, inadequate housing, and road accidents will become things of the past.

In London, the 120 m Millbank Tower, along the Embankment by the Tate Gallery, was built around a single reinforced concrete centre core. Centre Point, in St Giles' Circus in London, also about 120 m high, has two cores, at either end of the structure; the cores were constructed first and most of the building prefabricated at ground level and hoisted up into position.

A tower city will be built in the same way. People will be living and working in the lower part of the city long before its upper floors are finished.

The centre core of the 3 km high city will be like the trunk of a tree. Firm and tall, with foundations up to 400 m deep, it will accept the stresses passed on to it from its floors just as a tree trunk accepts the weight of branches.

Naturally, a great deal of detail work remains to be done. It would be misleading, for example, to claim that the linear lift system of three-dimensional personal transportation has achieved any degree of refinement. And the problems of waste disposal in a single building containing 300 000 people and their industries will obviously tax experts in that field.

In concentrating the activities of the community, the movement up and down within the tower would be very convenient, overcoming problems of external congestion. Improved communications, such as closed circuit television and vacuum-tube delivery, would eliminate a great deal of human traffic within the building. Leaving the lower levels as an open space amenity accessible within a few minutes would minimize the social effect of high concentration and skyscraper living.

There is also the question of public reaction to the prospects of living above the clouds. Today, some people are reluctant to live in comparatively small blocks of flats because they fear a structural collapse. Will they ever be able to accept the challenge of life in a city tower?

Simple arithmetic shows that the vertical city would make available for other purposes an area of well over 250 km² which in horizontal cities like Manchester and Bristol are occupied by buildings. Moreover, the city's transport problem would be almost eliminated.

Cost of such a development would be comparable with that of a modern 500 km motorway. Expenditure on such a scale would have to be weighed against the benefits the vertical city would bring. Of course, deeper investigation might well show that the cost would be appreciably less than that of providing a horizontal city of the traditional type. Certainly no country could benefit more than Britain from such a development. Science will create a new world and new way of living. Some people might even consider the venture worth embarking upon merely in order to live above the British weather.

Wilem W. Frischmann is a partner in the London firm of Pell Frischmann & Partners, consulting engineers. He has been responsible for some of the tallest buildings in the United Kingdom. He has lectured and written extensively on multi-storey structures and won several awards.

A. J. Kennedy

High temperature materials

Deterioration in materials at elevated temperatures imposes a fundamental barrier to progress in many important fields of technology. Deeper understanding of the solid state is combining with new techniques to push the barrier upwards.

Technological advancement is to a high degree linked with raising the operating temperatures of machines and sustaining these for prolonged periods. Usually the primary limitations are exerted by the properties of the available structural materials, and in the last 30 years a prodigious effort has been directed towards extending the capabilities of these so-called high temperature materials.

The term has little real meaning although by convention it is usually taken to apply to those materials capable of sustaining engineering stresses at temperatures in excess of about 800 °C or to those with melting points above about 1900 °C. But at higher temperatures there are many classes of material and several distinct temperature regions which need to be distinguished. At the same time a variety of materials operating at much lower temperatures could be regarded as fulfilling a high temperature role. For instance, aluminium alloys operating at 150 °C are subject to the same kind of considerations as those which enter into judgements on the behaviour of nickel alloys at 700 °C. Indeed, any metal exposed to a temperature in excess of about half its absolute melting point is vulnerable to thermally activated processes and particularly to time dependent deformation and fracture if the stresses are sufficiently high.

In today's advanced technological projects the position is made particularly difficult by the complexity of the operating conditions and by the interplay of a number of material factors each of which may exert a controlling influence on the behaviour of the whole system. It seems unlikely that unqualified empiricism can suffice any longer as a basis for judgement, if only because the testing programmes needed to establish the patterns of behaviour are uneconomic and unacceptably protracted. I shall not review these issues here but rather concentrate on some of the materials of immediate interest to high temperature technology.

Material properties

As the temperature is raised all metals show a characteristic loss of strength. A fairly critical change occurs between approximately 40 and 60 per cent of the absolute melting point. Creep becomes a significant factor here and so does the stability of the metallurgical structure – indeed the two may be closely related in a complex creep-resistant alloy. Diffusion-controlled processes all proceed at a significant rate in this band, so that profound changes may develop in the basic structure. The general relationship indicates the profit to be gained from using high melting point metals, although some base elements – titanium, for instance – lose strength at an unexpectedly low temperature. Moreover, the rule gives no indication of the improvements which might be realized by alloying, the nickel base alloys being outstanding in this respect.

But strength alone is much less significant than the ability to retain strength at high temperature, even if the absolute level is modest. Reaction with the environment, notably oxidation, is really the troublesome factor. Four metals with encouragingly high melting points – molybdenum, niobium, tantalum, and tungsten – are all vulnerable but are nevertheless the most important of the arbitrarily designated 'refractory metals'. Nearly all the important alloys in this group are formed by combinations of them.

Control of reactivity has proved much more difficult than control of mechanical strength. The latter is equivalent to increasing the difficulty of dislocation movement by governing the nature of the lattice or by the introduction of barriers. Both possibilities offer significant rewards and a large theoretical margin has yet to be utilized. Reaction rates of lattices of interest for high temperature applications are less affected by metallurgical variation, although the improvements attained by alloying and by the control of trace elements are welcome.

On the whole, the main weight of the research effort has been directed towards the achievement of protection by surface coating. This method looks fairly straightforward since some highly inert ceramic materials are capable of being formed into coatings. Unfortunately the very mechanical properties which make metals so valuable in engineering, particularly toughness or resistance to impact, are singularly absent in most of the thermally effective coatings. In addition this same distinction between the physical properties makes a matching of the two materials difficult. Under real operating conditions, and notably under severe thermal cycling, the mismatch may be so great that the coating breaks away from the underlying metal. One solution would be to make the whole component from the coating material, but the limitations imposed by the small amount of plastic deformation possible with such materials are severe. The component would be brittle and vulnerable to impact damage, and possibly difficult and expensive to fabricate or join to another. Nevertheless, in certain types of systems the advantages of the high strength-to-weight ratio of ceramics at high temperatures may prove over-riding.

Another fundamental factor is the extent to which an anisotropic structure may be developed and utilized beneficially. An example is a material for use as a protective layer under conditions of rapid heating to high temperatures for relatively short periods – exemplified by surfaces of space vehicles re-entering the Earth's atmosphere (see 'Atmospheric re-entry', *Science Journal,* June 1965). Much is to be gained in such instances by seeking to make thermal conductivity a maximum parallel to the surface and a minimum normal to it. The fact that such features may be developed in one material and not in another could have a large effect on the overall assessment. For example, it is a notable property of the highly oriented graphites formed by pyrolytic deposition, as distinct from the less oriented graphites made by conventional means, but the benefits are offset by certain features of the associated anisotropy in the mechanical properties.

The actual properties brought into play naturally have to be considered in the light of particular applications; the possible role of such properties as elastic modulus, thermal expansion and conductivity, yield point, and rate of oxidation are evident enough. But with the evolution of high temperature machines, resistance to deformation under a static stress component (creep) and resistance to fracture under an alternating stress component (fatigue) have become major and interconnected subjects. The type of stressing history imposed on machines almost inevitably incorporates both components, usually in a complicated way. Each exerts effects on the material structure in addition to thermal exposure itself. The elucidation of the long term behaviour of a material which is changing continuously in such conditions is basic to many current projects, but it remains a subject full of contradictions and uncertainties.

Economics

The task of summarizing the contemporary materials effort would be easier than it is if the various projects in which high temperature engineering plays a primary role had more in common. The demands for increased efficiency – whether in an aircraft engine, a nuclear reactor, or a steam plant – all bear on materials developments but the economics of each situation have to be brought into the reckoning if the real material incentives are to be properly assessed. In the case of electricity generation the replacement of half the carbon steel in the high pressure parts of boilers by alloy steel would increase the thermal efficiency from 25 per cent to between 40 and 50 per cent, but alloy steel is twice as expensive as carbon steel. Proper account has to be taken of all such economic facts and of the technical issues raised by ease of fabrication (particularly welding), by maintenance requirements, and by protection from corrosion, the last probably being the most persistent limitation.

The way in which high temperature factors enter into the electricity generating industry must necessarily differ markedly from their influence on space technology. The basic cost of material is over-riding in the case of most industrial products but in space technology it assumes a less exacting role (although still important enough) because a relatively small amount of highly specialized material may make the difference between success and failure in a project which is enormously expensive anyway. Moreover in space technology very high temperatures may be encountered as a necessary consequence of the mission so that it is not possible to design in more modest terms and fulfil the objectives of the mission.

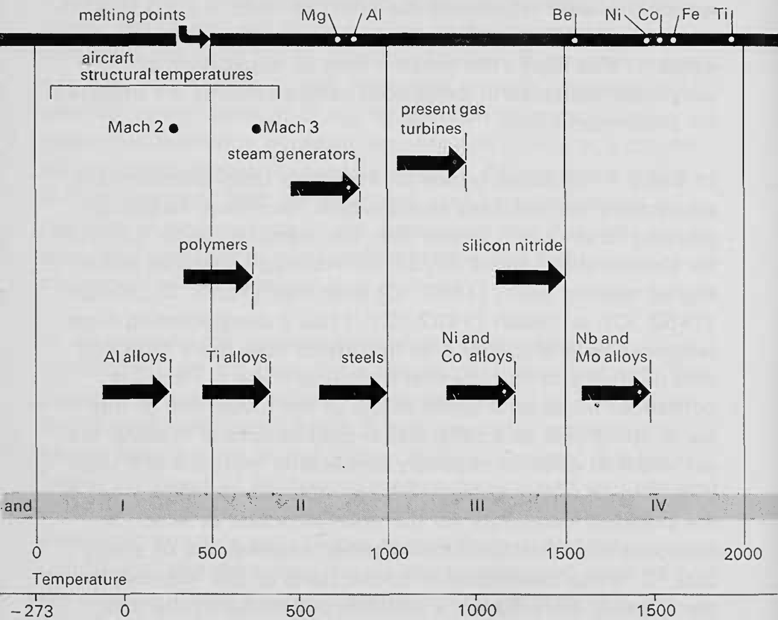
In aeronautics the development of the gas turbine has traditionally set the pace for materials evolution, but with the prospect of transport aircraft flying at Mach 2 or 3 – and no doubt eventually up to Mach 8 and beyond – materials for higher temperature power plants, such as ramjets, and for higher temperature air-frames have to be brought into use and evaluated. Nuclear engineering is similarly concerned on economic grounds with higher temperature operating conditions, and here new factors are encountered such as the effects of high temperature irradiation on materials and the resulting physical and chemical changes that may develop over prolonged periods of operation. At present such issues form a major subject in their own right.

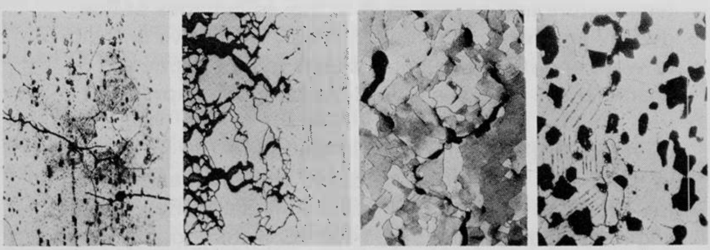
Useful temperature ranges

Each of the needs I have touched on above brings with it an associated set of economic and technical factors, and generalities are not easily derived. Before embarking on a discussion of the materials situation it might be worth while attempting to set out the temperature ranges of technological interest. I feel this can most clearly be done by accepting a rather arbitrary division into bands of 500 °C (figure 16). Normally the range 0–500 K (−273 to 227 °C) is not considered a true high temperature region but it nevertheless covers a critical thermal range for certain metals characterized by a transition from brittle to ductile behaviour. Conventional light alloys such as aluminium and magnesium cease to be practical engineering materials towards the upper boundary which probably represents the effective limit of such special variants as SAP, an aluminium matrix reinforced by dispersed oxide. It also marks the present limit of application of polymeric materials in conditions where stresses are imposed for prolonged times.

In Band II the alloys based on iron have been predominant, particularly the stainless steels which have been raised by alloying to the point where they can perform useful functions for thousands of hours at 750 °C. Although titanium has a higher melting point (1660 °C) than iron (1535 °C), nickel (1453 °C), or cobalt (1492 °C), it has a disappointing high temperature strength; but its density of little more than half that of steel and its chemical inertness make it a serious contender for special applications at the lower end of this band. Its virtues as a potential aircraft structural material are sufficient to offset a relatively low elastic modulus and high cost, and its alloys appear to have replaced stainless steel as the principal candidate for the main structure of Mach 3 transport aircraft encountering peak temperatures of about 300 °C. It has been used in cooler parts of gas turbines for many years. Beryllium is a possible contender at the low temperature end of this band. It has advantages of low density and high modulus but is very expensive and difficult to work and its low ductility seems to be an inherent characteristic which is unlikely to be much improved.

Development of engineering in the range 750 to 1000 °C in Band III is intimately associated with the gas turbine and the evolution of the nickel–chromium alloys. Raising peak engine temperature to, say, 1400 °C would double specific thrust but it is questionable whether the Ni–Cr series can be extended much further by conventional alloying. The cobalt alloys run fairly parallel in their behaviour but in general have not proved so useful, although the best of the cast alloys





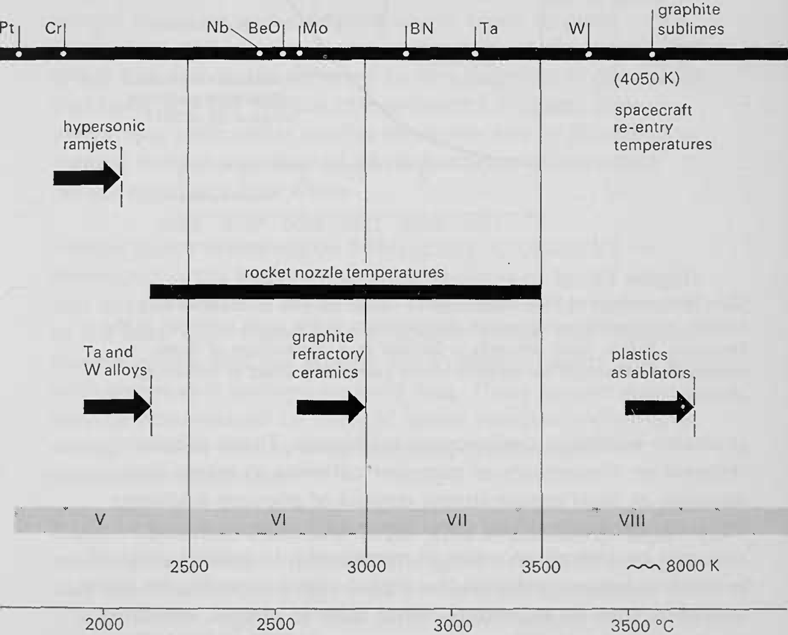
a
fatigue in high temperature aluminium alloy at 20 °C

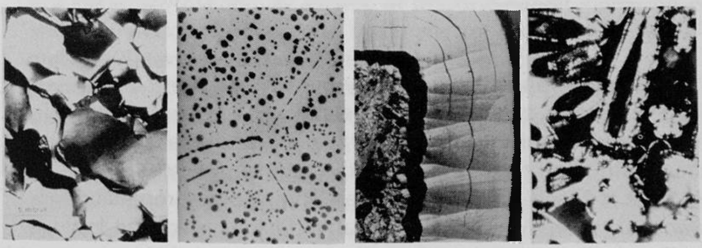
b
stress corrosion in nickel alloy at 650 °C

c
fatigue cracking in nickel at 816 °C

d
Cr–MgO composite exposed one hour in air at 1370 °C

Figure 16
The spectrum of temperatures of technological interest divided into bands each characterized by particular materials and their applications. The small photomicrographs of typical materials illustrate characteristics of the eight bands.





e
fracture surface of alumina at 1750 °C

f
internal melting in Nb–Ti alloy at 2245 °C

g
cracking in pyrolytic graphite coating

h
pyrolytic graphite/silica fibre matrix formed by surface reaction in reinforced epoxy resin

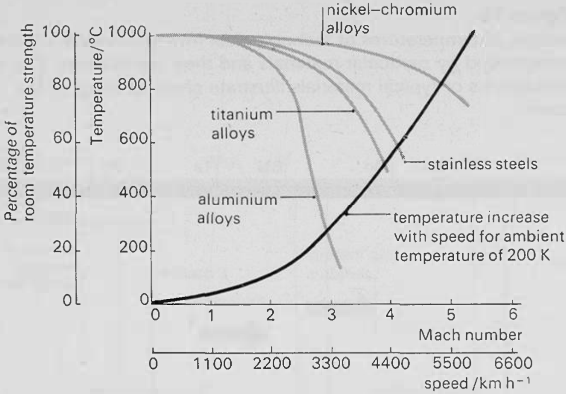


Figure 17
Skin temperature of high speed aircraft rises rapidly at supersonic speeds and causes progressive deterioration in the yield strength of the structure. Fall in yield strength is plotted as a percentage of room temperature strength for various alloys commonly used in advanced aircraft.

probably exhibit superior creep resistance. These alloys depend on dispersions of complex carbides to retain their strength at high temperatures; devoid of alloying additions (notably chromium) the time-dependent rupture strength is reduced by almost an order of magnitude. In certain parts of re-entering space vehicles the nickel alloys in particular come into their own as structural, rather than as engine, materials; but for space technology temperatures of 750 to 1000 °C are relatively modest.

In the Band IV range the refractory metals, and particularly those with body-centred structures, have special relevance. This coincides with a requirement in ramjets where the so-called stagnation temperatures in the diffuser inlet may reach this level at Mach 4 and above (see 'Hypersonic propulsion', *Science Journal,* May 1965). Suitable designs involving few moving parts and full use of cooling systems may keep temperatures close to 2000 K even up to Mach 8, remembering that flights at such speeds will be only tens of minutes in duration.

Of the 'short list' of refractory metal candidates molybdenum oxidizes catastrophically above 700 °C, niobium is better in this respect but has less desirable mechanical properties (a lower elastic modulus for example) while tantalum and tungsten, both higher melting point metals, impose heavier weight penalties with densities nearly twice as great. Fabrication and joining are problems with all the metals in this group but, taking the different factors together, it does seem that there is a fair chance of a protected niobium alloy performing sufficiently well at the lower end of this band to make a serious appraisal of its performance in advanced ramjet engines worth while.

Temperatures in the region 1750–2250 °C (Band V) are envisaged in the high temperature sections of future ramjets and already occur in the throats of smaller rocket nozzles. In the latter case materials are required which will withstand thermal shock and gas erosion for a relatively short time but with reasonably uniform material loss. These temperatures may be encountered by parts of space vehicles and missiles during re-entry and by magnetohydrodynamic power generation equipment.

Band VI between 2250 and 2750 °C is important in larger scale rocketry where materials are again needed to withstand very severe conditions for times of the order of minutes. It is a difficult regime to deal with, partly because of the very few materials which can survive at these temperatures anyway and partly because operating times are not so short that gross material deterioration can be tolerated. Graphite is outstanding in this respect, provided the system is not an oxidizing one, and has the added virtue of a density about one fifth that of steel. It has the interesting feature – in common with some other non-metallic materials, such as boron nitride – that its strength actually increases significantly with temperature over a limited range up to about 2700 °C; but in general it is too brittle to be used as a primary structural material. Between 2000 and 3000 °C significant plastic deformation can occur in graphite by creep. Studies at

Cranfield show the general features of this creep to be similar to those of metals at equivalent high temperatures.

The regime above 2750 °C covers combustion temperatures in high thrust rockets, and for convenience we might extend it upwards to include the very high but short term temperatures encountered by leading edges and nosecones of missiles and spacecraft – say 8000 to 10 000 °C. Resistance to such temperatures is achieved largely by ablative techniques involving the absorption of energy by the conversion of a solid to the vapour phase with some associated energy dissipation mechanisms such as combustion and, more important, radiation. Fortunately materials which fade out of consideration for long term usage at 2000 °C, such as reinforced plastics, now re-appear as highly useful for protective purposes at these extreme temperatures, where the duration of heating is short and material destruction is, within limits, acceptable.

Properties needed

For high temperature applications the classical engineering requirement of strength with adequate ductility is supplemented by that of low chemical reactivity. There is a certain perverseness in the fact that strength is often inversely related to ductility – which is just what would follow from a reduced dislocation mobility, other things being equal. In the same way chemical susceptibility to high temperature environment is inversely related to general mechanical toughness. Moreover refractory oxides are formed by such metals as aluminium, magnesium, and beryllium, which have low melting points; in the main, the oxides of the higher melting point metals are unstable and non-protective. The technological scene would be very different if an element existed with a melting point of 3000 °C capable of forming and maintaining an adherent protective oxide at a temperature as low as 1500 °C.

The extent to which the useful temperature range of the nickel alloys has been advanced, notably by chromium, titanium, and aluminium additions, is an encouraging indication of the possibilities that may well exist with the refractory metals. Alloys of this group already show two or three times the strength of the parent metal and significant further advances will certainly be achieved. But there is no question of the Band IV range being conquered as that below it has been. The nickel and cobalt alloys, with non-metallics such as silicon nitride and silicon carbide (both of which have limited possibilities because of their low toughness) probably represent the limiting front of single material application. Niobium alloys containing tungsten perform as well at 1300 °C as the best nickel-based alloy at 1000 °C, but only in an inert atmosphere. Progress beyond 1500 K requires special protection for these new materials, probably in association with dispersion hardening techniques. Both may be regarded as an aspect of composite material development and it is from this line of study that the main advances will be won in the future.

Protective coatings

Refractory coatings are an obvious means of protection but the attempts to produce and apply them have been discouraging. Some have good resistant properties in themselves – silicides, and in particular molybdenum disilicide with an alumina outer coating, or deposited silicon carbide, or even arc-plasma-sprayed coatings of refractory oxides and borides; but in every case serious limitations are encountered. With silicon carbide on molybdenum the bonding is poor. With molybdenum disilicide on niobium the coating embrittles the underlying metal although it has much the same thermal expansion coefficient and survivals for some hundred or so hours at 1400 °C have been reported.

Edges are always difficult, and coatings are prone to mechanical damage and to fissuring under thermal cycling. Resilience generally is hard to achieve. Surface plating with other alloys is a possibility and both conventional materials, such as nickel alloys, and less conventional ones, such as alloys of silicon–chromium and silicon–iron (which does provide an improvement) have been investigated. The use of successively applied layers, each designed to play a part in the

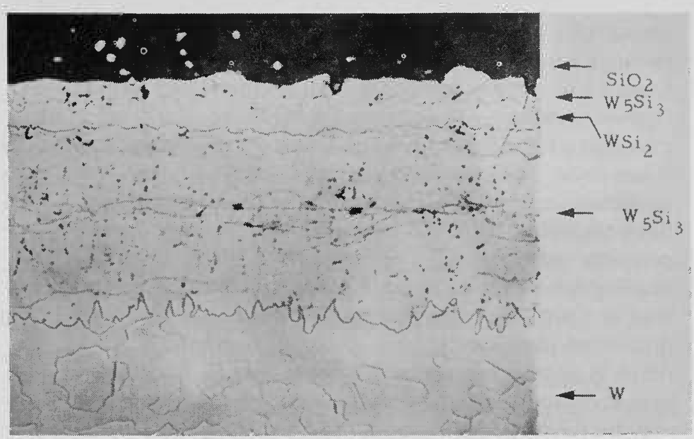


Figure 18
Layer formation at the surface of silicide-coated tungsten (W) after exposure to oxidation for 13 hours at 1600 °C. The various tungsten–silicon layers typify coatings generally.

overall compatibility and chemical stability of the surface region, is widely adopted. The simplest is probably represented by a diffusion-bonded silicide layer with a ceramic glaze overlay. The plating of one material on another can also produce a very effective improvement. Tungsten coatings formed from the decomposition of tungsten hexafluoride can be deposited on graphite, and tungsten can be plasma-sprayed on molybdenum preheated to a high temperature in nitrogen. In the latter case there is evidence of a true metallurgical bond, the substrate grains having dictated the nature of the surface layer growth and prevented any sharp boundary in either the structure or the mechanical properties – an important requirement.

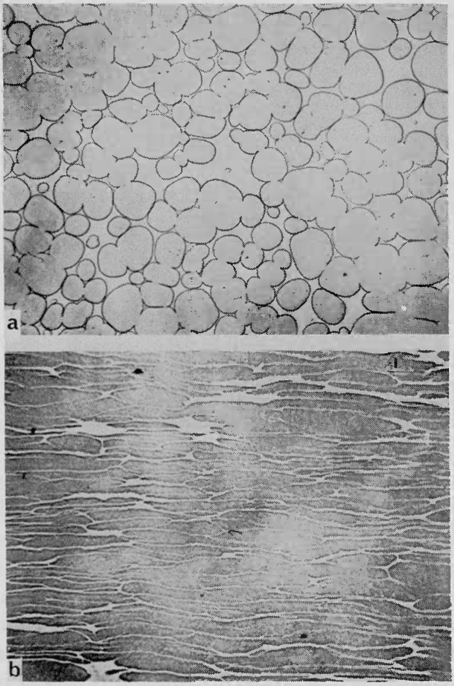


Figure 19

a Spheroidal particles of tungsten in a nickel–iron alloy. Polycrystalline tungsten is brittle but, when dispersed in this way, considerable reduction by cold rolling is possible.
b Micro-structure after reducing strip thickness by 90 per cent.

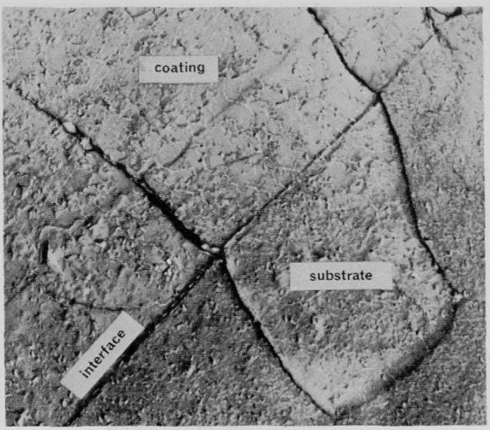


Figure 20
Plasma-sprayed coating of tungsten on a molybdenum substructure previously heated to 1350 °C. Close metallurgical bonding is evident; the lines are grain boundaries which continue in the coating.

New techniques

Such achievements underline the considerable role that new techniques play in materials development. The shaping of the harder high strength metals under hydrostatic pressure is one example. Another is the use of vacuum melting and the control of impurities, and a third is plasma spraying.

Returning to the basic material, high temperature developments in steels are of immense industrial importance to engineering projects with limiting temperatures of 600–650 °C. Examples are the evolution of families of chromium–nickel stainless steels containing relatively small amounts of the main four refractory metals, or the precipitation-hardening variety containing a little aluminium and well suited for building high speed aircraft. The martensitic stainless steels are particularly noteworthy in this respect: these are 10–12 per cent chromium steels, with say, molybdenum, niobium, and possibly cobalt additions. They are easy to fabricate and retain at least one third of room temperature strength at 650 °C. Though promising, chromium raises a number of difficulties as a basic material. It is handicapped by low temperature brittleness, although some of the alloys formed with refractory metal additions might possibly prove valuable within an operating range of 1100–1200 °C.

Dispersed-particle hardening remains an attractive possibility. The attempted utilization of two or more materials bonded together in a composite form is a feature of the contemporary materials scene. The addition of appropriate refractory particles – as exemplified by magnesia in chromium, alumina in aluminium, or thoria in nickel – can certainly improve high temperature properties but at the expense of strength at lower temperatures. The boundary conditions are critical and perhaps not yet properly understood, but in general the presence of the particles appears to impede the progress of those dislocation processes responsible for high temperature softening and creep. At lower temperatures, where such stress relaxation processes cannot occur, the particles may act as stress raisers and so lower the bulk strength of the aggregate. This is not an inevitable consequence and with certain kinds of dispersion (for example, precipitates) low temperature strengthening also results.

Although there is good evidence that the addition of suitable fine particles can improve high temperature properties the geometrical factors may not be particularly simple. For example, SAP (see page 49) retains a fair strength at temperatures even higher than the melting point of pure aluminium. There is considerable anisotropy in the deformation behaviour of the extruded material, suggesting that the continuous phase elongated alumina cell structure is more controlling in its action on high temperature deformation than are the dispersed particles. The particles nevertheless have a major effect on work-hardening characteristics and result in high dislocation densities and, ultimately, higher deformation stresses.

A number of variations on this basic theme have been attempted. In some cases metals may be dispersion-hardened by chemical means, as distinct from mechanical mixing, by the addition of particular constituents and by their subsequent precipitation – so called internal oxidation is a case in point. Significant room temperature increases in the proof stress for systems such as copper–alumina, nickel–alumina, and iron–alumina have been reported. For example, the proof stress of iron is doubled by two per cent addition of alumina.

Chromium reinforced by five per cent magnesia is now a commercial proposition. It is not as good on a strength-to-weight basis as a coated refractory metal such as a molybdenum alloy but is probably more reliable in service. It has a low rate of reactivity due, in part, to the formation of an adherent coating of oxide. Two per cent addition of tantalum carbide to chromium improves the creep resistance, largely by raising the recrystallization temperature as well as by a particle dispersion effect. But coating the edges of refractory metals is extremely difficult. A more likely contender for high temperature applications is nickel reinforced by thoria (TD nickel). Its 100 hour rupture strength and yield strength as a function of temperature are shown in comparison with other high temperature materials in figure 21. Ductility of this particular material actually declines with increasing temperature because of a critical grain boundary failure; in general its ductility tends to restrict fabrication. Rate of loss of material at temperatures above 1000 °C is sufficiently high to curtail significantly the use of thin sheets. Coatings based on aluminium or chromium are helpful but under thermal cycling these can fail by interdiffusion and void formation. Cladding with a nickel alloy gives better resistance. Joining involving melting is difficult and is a recurring problem with composites generally; but brazing, using palladium alloys, is possible.

Development of successful fibre-reinforced materials also has high temperature relevance. There is a scarcity of experimental data in this subject, partly because of commercial caution

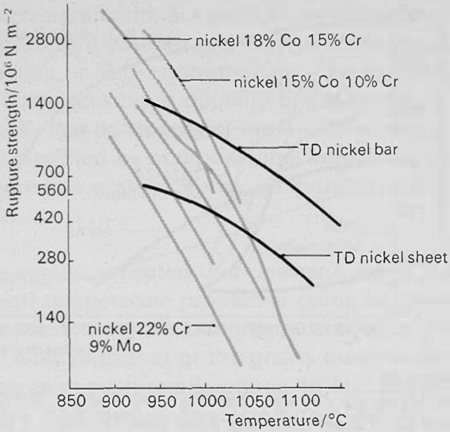


Figure 21
Thoria dispersion-hardened nickel falls off in strength much less rapidly with increasing temperature than do ordinary nickel alloys. TD bar and sheet curves are printed in black.

but mainly because the reinforcement of metals by this means has not yet been carried very far. Results of experiments on silica-reinforced aluminium are reproduced graphically in figure 22, showing the considerable strength gain over both SAP type material and an aluminium alloy. Metal wires can be produced with very high strength and their aggregation into a matrix of another (preferably chemically protective) metal is readily possible technologically. Boron fibres, with strengths approximately double that of the strongest steels and with a density less than aluminium, are particularly promising. But the boundary between the two metals is mechanically and chemically vulnerable, and the actual engineering utilization of these highly anisotropic systems raises some serious difficulties. Moreover, joining methods are an even bigger problem than with dispersion-hardened materials, for a number of reasons, some of which are self evident.

Some engineering projects call for just as much effort to be applied to raising the working temperatures of non-metallic materials. This may be because their unique properties, such as

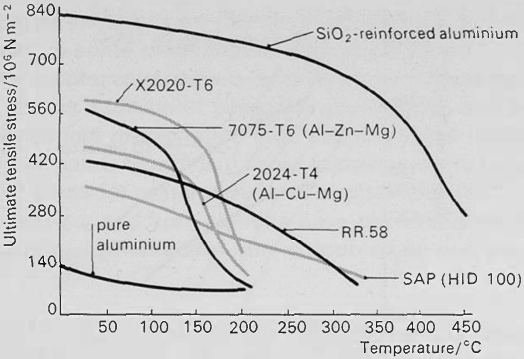


Figure 22
Aluminium alloys are stronger than the pure element. RR.58, containing 2.5 per cent Cu, 1.5 per cent Mg, 1 per cent Fe, and 1.2 per cent Ni, behaves well up to 200 °C; SAP, with 10 per cent oxide, overtakes it at 300 °C, but aluminium reinforced by silica fibres is overwhelmingly better if its inherent disadvantages can be overcome.

electrical or thermal insulation or flexibility or transparency, meet some highly specialized need. The polymers containing fluorine are of particular interest in this connection. Polytetrafluoroethylene, otherwise known as PTFE or Teflon, offers high strengths up to about 300 °C and variants on the basic carbon–fluorine bond structure look promising as both plastic and elastomeric materials for high temperature use where chemical inertness is a valuable feature – as it is in fluid sealing. The Viton A type, a copolymer of vinyl fluoride with hexafluoropropylene is an example of such a material in the elastomer category.

Future developments

For the future there is no single panacea so far as materials are concerned. A great variety of high temperature materials are required and each must be studied and assessed in its own context. Briefly the probable future course of evolution can be summarized as follows.

There are real strength margins which could be realized in metals by compositional and thermal control, particularly for low temperature application. At higher temperatures the benefits, though significant, are unlikely to bring about a serious 'breakthrough' without recourse to other measures such as dispersion or fibre strengthening. The promising materials are either lacking in ductility or too prone to oxidation attack. The brittleness of most of the ceramics is unlikely to be modified by compositional variations, but materials of very fine grain size may be more successful in this respect.

It has been generally accepted that coarser grained materials have better high temperature properties (such as creep resistance) because the grain boundaries are regions which lose strength, relative to that of the grains themselves, to a significant degree at such temperatures. In fact this is not necessarily true, and there is good evidence that the boundaries can actually be strengthening agents at high temperatures. Plotting a creep rate parameter against grain diameter for various metals reveals creep rates rising with increasing grain diameter in an approximately linear manner.

We should not overlook the more general possibility that new techniques, such as shock deformation, may enable substructures to be achieved which are not attainable by other means. Fine crystal grain size is a case in point. These developments will influence ceramics and metals alike. Ceramic coated metallic structures, often of complex form, such as honeycomb sections, are important high temperature engineering developments. The additional fibre reinforcement of such coatings is yet another line of variation. The emphasis generally is shifting towards using composites of this kind in which the inherent limitations of the individual materials are accepted but in which each has a vital role to play.

More and more one recognizes the dominant part that interfaces and boundaries play in such a situation, and fundamentally it is this field of study which is bound to come forward strongly in the next few years. We have spent an enormous effort, with generally rewarding results, in elucidating and controlling the strengths of crystal lattices, but far less on the interfacial region between two dissimilar lattices. This is a region which is responsible for many of the current high temperature difficulties. At the same time radical changes are required in attitudes to design. The special characteristics of the new materials must be better taken into account, not merely in calculating their probable behaviour in the engineering system itself but in assessing the possibilities of their economic fabrication into acceptable forms.

Just as materials science is breaking down some of the traditional subject boundaries, so must materials engineering break through the established divisions, widening the scope of design as we now know it and utilizing more effectively the knowledge we have. This kind of revolution is necessary if the problems encountered in high temperature engineering are to be handled rationally, and we may eventually see these changes as one of its most valuable by-products.

Dr Alfred James Kennedy is Director of the British Non-Ferrous Metals Research Association and was previously Professor of Materials at the College of Aeronautics, Cranfield, Bedfordshire. Before that he was Head of the Metal Physics section of BISRA, Royal Society Armourers and Braziers' Research Fellow (at the Royal Institution), and a physics lecturer at London University.

Acknowledgements

Figure 16: (a) College of Aeronautics, Cranfield, Bedfordshire; (b) Douglas, D. A. (1959) *High temperature materials,* John Wiley; (c) Achter, M. R., Danek, G. J., and Smith, H. H. (December 1963) *Trans. Met. Soc. AIME* **227**; (d) Manning, C. R. and Royster, D. M. (1963) *Technical note D-1785,* NASA; (e) King, A. G. (1961) from paper in *Mechanical properties of engineering ceramics* (page 340), Interscience; (f) Rostoker, W. and Dvorak, J. R. (1965) *Interpretation of metallographic structures,* Academic Press; (g) Diefendorf, R. J. and Stover, E. R. (May 1962) *Metal progress,* **81**, 108; (h) Walker, A. C. and Scala, E. (1963) from paper in *Materials in space technology* (chapter 11), The British Interplanetary Society.

Figure 18: Klopp, W. D. (1962) *Report 167,* Defense metals information center, Battelle Memorial Institute, Columbus, Ohio.

Figure 19: Rostoker, W. and Dvorak, J. R. (1965) *Interpretation of metallographic structures,* Academic Press.

Figure 20: Spitzig, W. A. and Grisaffe, S. F. (1964) *Technical note D-2510,* NASA.

E. R. Laithwaite

New forms of electric motor

Fundamentally novel types of electrical machine have recently evolved in a search for concepts able to overcome the handicaps of traditional designs. Research in this field has acquired a new spirit of adventure.

The first electric motors were developed by some of the greatest inventors of all time. They produced not only economically sound designs but also analogies which made the principles of their machines readily understood. Yet, as is often the case with concepts which are ahead of their time, subsequent exploitation of electromagnetism in machines was less adventurous.

There are fashions in engineering just as there are in clothing. Like the tailor, the machine designer seeks to please his customer and, in the same way that the farm worker will not want the same kind of cloth for his working trousers as the lawyer does for his, different applications of electrical machines will emphasize different properties. In modern machines power-to-weight ratio is more important than it was 40 years ago when efficiency and power factor were the fashionable properties. Since 1945 there has been an increasing awareness among engineers that machines can often perform a useful purpose even though their characteristics do not conform to orthodox standards.

Such reappraisal of the usefulness of electrical machines has now opened the way for a re-examination of past inventions which had been discarded on the grounds of low efficiency or high cost. As engineering ventures become more ambitious the demands on ingenuity become greater and the question is not so much 'How efficiently can it be done?' as 'Can it be done at all?' Machines have now been developed with topological structure basically different from the conventional layout. New types of solid state device, such as the silicon controlled rectifier (thyristor), have changed the fashion in variable speed drives and made the direct current (d.c.) machine even more popular, but it is proposed to review not new auxiliary apparatus but machines which are structurally new.

Principles of d.c. motors

Electrical machines are all based on the fact that force is produced on the moving part by the reaction between a magnetic field or flux and an electric current. Every electrical machine is able to reverse its energy conversion process: a motor converts electrical energy into mechanical energy but it can equally well function as a generator and convert mechanical energy into electricity. I propose to relate this article to motors, but it must be remembered that the arguments apply to generators also.

In the simple d.c. machine sketched in Figure 23 a coil of wire is free to rotate between the poles of a magnet. When current is passed through this rotor coil in the direction of the black arrows, forces will be produced acting on the rotor in the direction of the tinted arrows. These forces form a couple which will turn the rotor. When the rotor has turned through

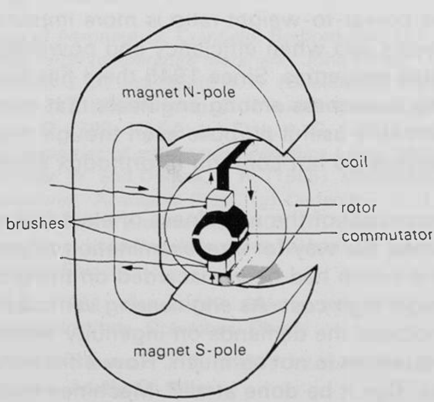


Figure 23
The conventional d.c. motor.

90° the original forces on it will have fallen to zero. Further rotation will make the rotor coil come under the influence of the opposite poles and the forces on it will be reversed.
To produce continuous rotation the current in the coil has to be reversed as the coil passes the horizontal position. This is done by the commutator, a rotary switch inserted between the input terminals and the coil. In a practical motor the rotor usually contains many coils each connected to its own pair of commutator segments.

Whereas direct current is equivalent to steady charge movement, alternating current (a.c.) is produced by accelerating charges and has a sinusoidal wave form. This makes possible the transfer of energy from one circuit to another by transformer action; the changing current in one coil can induce a changing current in another without the need for physical contact. In an a.c. machine the rotor current is induced in this way so that no commutator or brushes are needed. There are three basic types of a.c. machine: induction, synchronous, and commutator.

Basic a.c. motors

In the induction motor a.c. is fed to a series of coils each energizing one of an even whole number of magnetic poles arranged round the stationary member or stator. The resulting stator currents generate a rotating magnetic field, the speed of which is such that it progresses one wavelength, from one north pole to the next, for each cycle of the supply. The rotational speed is thus given by the ratio of the supply frequency to the number of wavelengths or pole pairs. Inside this field is placed the rotor which may comprise either a series of closed electrical circuits set in an iron cylinder or merely a solid drum of copper or aluminium (see figure 24). The rotating magnetic field generates secondary currents in the rotor by transformer action, and the action of the rotating magnetic field on these currents is such that it tries to drag the rotor round with it. Unless there is relative motion between the field and the rotor there can be no rotor currents and no force. For this reason an induction motor can never run quite as fast as its field.

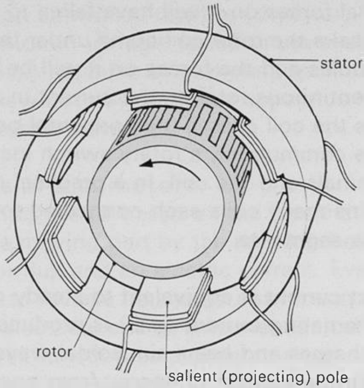


Figure 24
Induction motor. (The tint on the rotor shows one arrangement of conductive paths.)

In the synchronous motor the stator is virtually the same but the rotor consists of a magnet which tries to align itself with the rotating field. If it succeeds the motor will run at precisely the speed of the field. This is the only possible running condition, and synchronous motors are therefore used in clocks. Increasing the torque load makes the rotor lag slightly behind the field until eventually the motor stalls. The third type of a.c. machine, the a.c. commutator motor, is effectively a combination of an induction motor with a d.c. machine.

The induction motor is by far the commonest form of electric drive. It has been estimated that it accounts for more than 95 per cent of the electric horsepower in the world. Since its moving part can consist of a solid piece of metal it is inherently robust and reliable. Its disadvantages are that it is basically a constant speed machine and that, if its efficiency is to be high, the peripheral speed of its rotor must be at least of the order of 6 m s^{-1}. Modifications to produce a variable speed induction motor have usually involved introduction of brushes and rubbing contacts, involving sacrifices in power-to-weight ratio, reliability, and cost.

One notable exception existed in the pre-war period. It was recognized that, if the connections to half the stator coils of an induction motor were reversed, the speed of the field could be doubled. It was therefore possible to produce a highly efficient two-speed machine using only one stator winding, with consequent saving in copper, but the speed ratio had to be 2:1.

New developments in multi-speed induction machines began simultaneously in 1955 at the Universities of Bristol and Sheffield, with the development of single-winding motors capable of two speeds in the ratio 3:1. The engineers developing these machines had the advantage of a background of light-current knowledge, and radio techniques clearly influenced the important developments which followed.

In 1958 Professor G. H. Rawcliffe and his co-workers at Bristol introduced the idea of 'pole amplitude modulation' for induction machines. Although an induction motor rotor must contain an even whole number of poles it is clearly possible to change the field speed, towards which the rotor is accelerated, by so altering the flux configuration as to change the effective number of poles.

This effect can be achieved by the radio technique of multiplying one wave by another. For simplicity of illustration this can consist of merely reversing the connections to one half of the winding, for this can be said to multiply the original wave by a square wave. As shown in figure 25, such multiplication leads to a wave which approximates closely to the sum of two constant amplitude waves of different wavelength. The invention of the new machines came with the realization that, in a polyphase system, the phase windings can be so spaced that one of the two additive waves can be eliminated entirely without detracting from the value of the other. This original breakthrough led quickly to a series of further inventions by the Bristol team, extending the range of speeds which could be obtained, increasing the number of possible speeds in a single machine, and reducing the harmonics which result from the fact that a limited number of stator coils cannot achieve perfectly sinusoidal flux distribution.

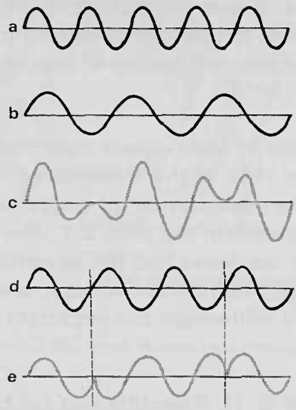


Figure 25
The ten-pole wave (a) and the six-pole wave (b) can be added to produce wave (c); if an eight-pole wave (d) has the portion between the broken lines inverted it produces (e), which behaves very much like (c). Thus inversion of part of the waveform has the effect of changing the number of poles and hence the speed of the motor; this is the basis of pole amplitude modulation.

Rawcliffe pole change induction motors are an example of ingenuity in arranging a machine's electric circuit. Every electromagnetic machine can be said to consist of a magnetic circuit and an electric circuit interlinked. The basic difference between these circuits is that, whilst electrical insulators are capable of confining the current to the conductors, there is no equivalent magnetic insulator which will prevent magnetic leakage. The techniques of design are therefore different.

Linear motors

Whilst the work in Bristol was proceeding other research at Manchester University was concentrating on developing linear motors – a family of machines which essentially involve a rearrangement of the magnetic circuit. Linear machines are, in effect, the physical embodiment of a 'developed diagram', resembling a conventional cylindrical machine cut along a radial plane and unrolled to produce a flat strip motor capable of propelling its rotor in a straight line (the word 'rotor' is retained for linear machines because no ambiguity usually arises and a new word seems unnecessary). At first sight this results in no fundamental change in principle, yet three new operating features must be considered.

The machine now has 'ends' in the direction of motion and, as soon as the rotor moves, a part of it slides off the stator at one end and exposes a corresponding length of stator at the other. It is therefore generally necessary to elongate one member or the other to make either a 'short stator' or 'short rotor' machine. Both machines waste material in that a portion of them is inactive. The second point is that the ends of the shorter member distort the flux pattern and so may reduce the power-to-weight ratio and increase the iron loss; and, when the rotor moves, transient currents set up by the moving edges produce excess resistive losses in the windings.

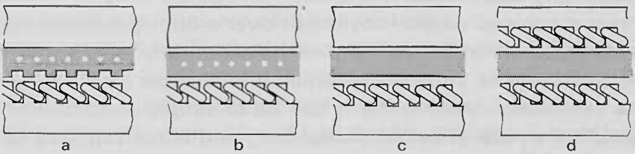


Figure 26
A linear motor can be evolved by opening out a conventional induction motor (sketch (a) shows a small portion with rotor above and stator below), severing the rotor conductors (tinted) from the original rotor iron and linking the latter to the stator (b), simplifying the rotor to a solid sheet (c), and finally adding a winding on the other side of the rotor (d). In practical transport or machine tool applications the rotor would normally be fixed and the moving part would be the iron and windings on either side of it.

Thirdly, in addition to the induction force between rotor and stator members, which may be likened to a shear stress, there exists a purely magnetic compressive force between the two (since both contain magnetized iron) which can be even greater. In a rotary machine the magnetic forces on the rotor cancel each other everywhere except for a small residual pull due to mechanical imperfection. In a linear machine the magnetic pull is entirely on one side. The designer of linear machines is always struggling to minimize these unwanted effects.

Applications of linear motors were, for a long time, thought to be limited chiefly by the fact that there is no convenient linear equivalent of the gear. Most linear motions required by industry are at speeds less than 3 m s^{-1} and induction action is not particularly advantageous at such speeds if the user is concerned with such quantities as efficiency and power factor. Simplicity of action, a high degree of reliability, low maintenance costs, flexibility of application, and absence of noise however are all features of linear machines which have persuaded engineers in recent years to accept low speed linear motors or 'actuators' despite possible low values of efficiency and power factor.

Linear motors may be divided into classes in a number of ways. One classification is shown in figure 27. In linear machines where movement over a considerable distance is involved either the primary or the secondary member must be elongated as shown. Generally it is cheaper to elongate the secondary member, for it can be of simple construction (such as a plain sheet of aluminium) and is not required to be fed with current directly.

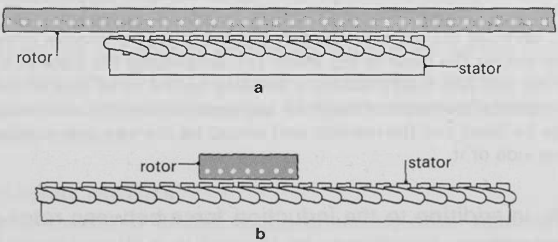


Figure 27
Classes of linear machine.
a Short-stator machine.
b Short-rotor machine.

Another way to divide linear motors is by application, in which case there are three main classes:

1 Power machines – usually high speed, high efficiency motors.

2 Energy machines – principally used to *accelerate* masses over short distances.

3 Force machines – for low speed and standstill applications.

The last two of these have already been commercially exploited, particularly the third. Linear thrust units are now used to operate travelling cranes, sliding doors, parcel sorters, conveyor belts, X-ray plant, etc. A large Anglo–French organization 'Lintrol Systems' now manufactures linear motors in quantity for these purposes. Two large accelerators have been in use for a number of years: one installed at the Motor Industry Research Association Laboratories at Nuneaton is used for crash-testing road vehicles; the other at the National Engineering Laboratory at East Kilbride is used for rope testing.

High speed machines are about to be tested in Cambridgeshire by Tracked Hovercraft Ltd as drives for high speed vehicles (400 km h^{-1}) travelling on air cushions. Similar work is being carried out in the United States, in France, and in Japan. All will use the linear motor which has the advantages over other forms of drive, such as the airscrew and rocket, that it neither pollutes nor disturbs the air, is silent, and has no moving parts.

Future developments

Attempting to forecast possible technological developments is always hazardous and seldom accurate. Nevertheless it does appear that one new theme is emerging in electrical machine research: concentration not on the electric circuit, which has been the main subject of attention over the first half of this century, but on the magnetic. Except for certain special applications, notably those for aircraft, magnetic circuits of electrical machines remained virtually unchanged from 1900 to 1950. Basically the magnetic circuit of a conventional machine is two dimensional, the axial third dimension being merely adjusted according to the desired power output. Modern aircraft generators make more use of the third dimension, in that axial flux is used to simplify the construction and make the moving member more robust for operation at high rotational speeds.

In 1965, two new machines have been described which are the result of research devoted entirely to modification of the magnetic circuit. The first, disclosed in two papers by P. J. Lawrenson and L. A. Agu of Leeds University, is the reluctance motor in which a rotor containing no windings is made to lock with a rotating field by virtue of saliency (protuberance). Conventional machines of the reluctance type have used rotors of a shape which at first sight might appear to be the best. Yet the Leeds workers have shown how topological ingenuity can produce a better shape (see figure 28).

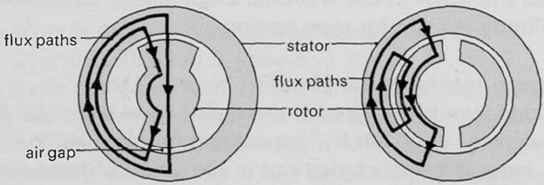


Figure 28
The reluctance motor modified by Lawrenson and Agu (righthand sketch) has reduced rotor inertia and better torque and power factor, all achieved by a new shape of rotor giving improved flux paths.

The second invention may be even more significant, for it comes from the Bristol group which hitherto had been concerned essentially with electric circuits. In their new machine the highly successful Rawcliffe pole change windings are used in combination with a salient pole rotor of such a shape that it is capable of accepting either of the pole numbers set up by the stator windings. By means of this ingenious magnetic circuit W. Fong and Rawcliffe have produced a two-speed synchronous machine to complement their induction or asynchronous pole change motors.

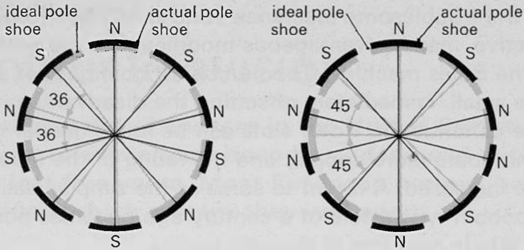


Figure 29
The pole shoe arrangement by Fong, French, and Rawcliffe can accept ten- or eight-pole flux. Pole shoe positions in conventional machines are shown in tint for a ten-pole (left) and eight-pole machine. The compromise configuration of rotor can thus be made to run at either of two speeds.

I will make one exception to my intention of not dealing with developments in thyristor controlled machines. The work at Imperial College of Dr J. J. Bates in 1963 involved radical changes in machine construction and, more important, represents entirely new thinking in regard to commutator machine operation.

The commutator of the d.c. machine has been a commercially acceptable device for many years, but it has limitations and disadvantages and the advent of the thyristor brought a hope that commutators might be eliminated. So far this has not been realized, but there is no reason why the best features of both mechanical commutators and solid state devices should not be combined in one machine. Dr Bates' machine carries a double commutator with brushes fed from thyristors in such a way that the brushes are never called upon to make or break the current. Instead, the current is switched by a thyristor when the corresponding brush is entirely contained by a commutator segment. When the brush crosses to the next segment it is electrically inactive and its task is reduced to that of a brush in contact with a slip ring. The number of thyristors required is only four, and their voltage rating is that of the voltage induced in the commutating coil by the interpole, and is a fraction of the machine voltage.

Once the troublesome reactance voltage has been made ineffective, many advantageous modifications are possible with the Bates machine. The number of commutator segments can be small, immediately offsetting the disadvantage of the double commutator. Rotor slots can be made deeper to accommodate more copper, and the rating of the machine can be increased. A return to some of the simpler, salient pole/bobbin type rotors of a century ago becomes possible and a whole new line of thought begins.

It is such ideas which are the more exciting feature of present research in electrical machines. Engineers with knowledge of both light and heavy current techniques are now looking at machine problems with an open mind – with well established structures temporarily pushed aside, lest they give the impression that to do something radically different is essentially a retrograde step. What is perhaps even more important is that their work is demonstrating to the present generation of students that not all the excitement has yet gone from the field of heavy electrical engineering.

Professor Eric Roberts Laithwaite is Professor of Heavy Electrical Engineering at Imperial College of Science and Technology, London. He was previously a Senior Lecturer at Manchester University and, earlier, an RAF officer at Farnborough. His current research embraces linear induction motors, electromagnetic levitation, and induction actuators and accelerators of various configurations.

F. C. McLean

Colour television

Already a noted success in the United States, colour television has become widespread within the last five years. Most European nations have decided which system they will adopt.

Regular transmission of colour television started in the United States in 1954. It made very slow progress for the first ten years and few receivers were in use. It then started moving rapidly so that by 1970 the greater number of television receivers sold in that country were for colour. In Europe regular transmission of colour started first of all in the United Kingdom in July 1967 followed very quickly by colour in Germany and in France and later in many other countries in Europe. In Europe the rate of growth of colour television has been much more rapid and by the early seventies the sales of colour receivers in the United Kingdom approached the number of black-and-white receivers sold and exceeded them in value. Transmissions were almost wholly in colour.

The United States colour television system is known as the NTSC (National Television Standards Committee). In Europe the situation is much more complicated: three separate systems were studied. The first was the American NTSC. (adapted to 625 line standard). The second known as PAL (Phase Alternating Line), was developed in Germany but closely resembles the NTSC and is widely favoured. The third, SECAM (Sequential Colour with Memory), system differs further and is favoured by France and some other countries. I shall discuss the respective merits of the three systems later in this article.

Following on a recommendation of the Television Advisory Committee, the British Government decided to accept the PAL system on 625 lines. This system is in use in a very large number of countries in Europe and elsewhere in the world.

Basic process

Colour television is in many ways very similar to colour printing. In colour printing three separate pictures of selected colours are superimposed, one after the other, to give the required pictures. In colour television pictures in three colours are superimposed simultaneously. But, whereas colour printing uses reflected light and the colours chosen are three selected 'primaries' of red, yellow, and blue, colour television works with direct light and the three colours required are red, green, and blue. Accordingly, what has to be done in colour television is to find a means of analysing a picture into its red, green, and blue components, transmitting these efficiently and economically, and then recombining them in the viewer's home.

The ways in which this can be done are limited by some important restrictions. The first of these is concerned with availability of channels. Radio frequency channels for television broadcasting are very scarce and it is essential that the minimum demands should be made of them. Secondly, colour television will initially be viewed by only a very few people in any area changing over to colour, and it is necessary to ensure that the colour transmissions can also be received by people equipped only with black-and-white receivers. Conversely, initial colour transmissions will form only a portion of the whole transmissions and the systems must be such that black-and-white pictures can be sent and be adequately received on colour receivers. The combined requirement that colour should be received on a black-and-white receiver and that black-and-white pictures should be sent over the same chain is called 'compatibility' and is absolutely essential. Thirdly, the method chosen for the transmission of the pictures must be economic at the studio and transmitter and, in particular, at the receiver. Economics carries with it the requirement of reliability, and the more components there are in a piece of equipment the less reliable it is.

As a result, the essential basis of all the possible systems is to send a signal identical with that used in a normal black-and-white transmission and to add to this the minimum of information sufficient to enable the receiver to regenerate the three colour signals. The black-and-white transmission is determined by the 'luminance' signal, and the colour content of the picture by the 'chrominance' signal. The problem is therefore: to analyse the picture to be sent into the red, green, and blue components and obtain the corresponding signals; to code these signals in a suitable form with separate luminance and chrominance signals; to transmit the coded signals; to receive the coded signals and decode them into separate red, green, and blue signals; to display the picture corresponding to the individual red, green, and blue; and to do all this in such a way that there is no interference to normal black-and-white reception. The basic process of forming a television picture by a line-by-line scanning process is, of course, the same in colour as in black-and-white transmissions.

White light is a mixture of all the colours of the spectrum but if appropriate shades of red, green, and blue are combined in suitable proportions the result looks like white light, while combinations of any two produce other colours. Almost any colour can be produced by mixing appropriate amounts of the three primaries. The human eye cannot distinguish between light consisting of radiation in any given narrow band of visible wavelengths and light that consists of mixed radiation in widely separated bands. For example, yellow light with a wavelength of about 600 nm is indistinguishable from an appropriate mixture of red near 700 nm and green around 530 nm. This is a great help. The luminance signal conveying the information for a black-and-white picture is electrically equivalent to the sum of the signals representing the red, green, and blue. If the chrominance signal is made to contain two signals electrically equivalent to the red and green then, by separating these out in the receiver, the red and green are obtained individually and, by subtracting the sum of these two from the white signal, the blue is obtained. This is the essence of all proposed methods of achieving colour television; although there are variations in the actual content of the chrominance signal, and in the way in which it is transmitted, basically all are similar. We therefore have to analyse the white light to obtain the signal in the three colours and then build up the luminance and chrominance signals to the form required by the chosen system.

Analysis of the white light is done with dichroic mirrors. These consist of a glass base on which are deposited from 7 to 20 very thin layers of transparent material. Each layer has a thickness of only a fraction of a wavelength of visible light and, according to its individual design and deposition, transmits certain colours and reflects others. By the use of a number of layers fairly steep changes of reflectivity as a function of wavelength can be obtained. In figure 30 the white light source is shown as a mixture of red, green, and blue. When this meets the first dichroic mirror the blue light is reflected and the green and red pass straight through without reflection or appreciable absorption. The green/red then strikes the second dichroic mirror which has a coating which reflects the red light and lets the green light pass through.

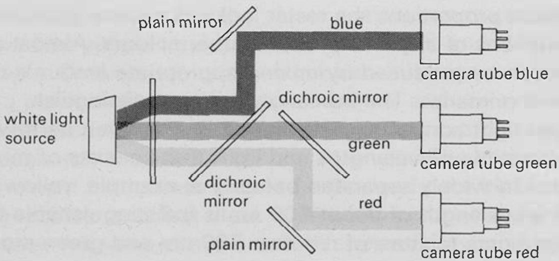


Figure 30
A dichroic mirror system is used to analyse the white light from the image into three colours ready for transmission. The mirror system is normally built into the television camera.

At each of the three outputs of the dichroic mirror system is placed a television camera tube employing normal methods of scanning the scene. By a combination of electrical and optical adjustments the pictures produced in the three camera tubes are made identical in size, providing three otherwise identical pictures giving the red, green, and blue content of the scene. The three colour signals pass through the chain to the receiver in which the three colour pictures are accurately superimposed. Thus the original picture is reconstructed.

Colour transmission

The colour television process is started by building into a television camera a colour analyser, using either dichroic mirrors or some other device derived from this principle, and providing three separate camera tubes. From the latter the three colour outputs, adjusted to give pictures of identical dimensions, are added together to give the overall black-and-white brightness: this is the luminance or 'Y' signal. The luminance signal is used to modulate a carrier exactly as is done in black-and-white television, thus ensuring the compatibility of the system. A proportion of the red, green, and blue outputs is used to derive the two signals required to give the red and green, or whatever other combination has been chosen, in the chrominance signal.

An analysis of the energy distribution of a television signal (figure 31) shows that all the information is contained in fairly narrow bandwidths centred on multiples of the line frequency. There is comparatively little information between these discrete frequencies, and comparatively wide frequency spaces centred on multiples of half the line frequency contain little information. Points of minimum energy occur all along the scale. The chrominance information is therefore packed into the spaces between the information in the luminance signal so as to cause a minimum of mutual interference; an adequate chrominance signal can therefore be sent within the bandwidth occupied by the luminance channel.

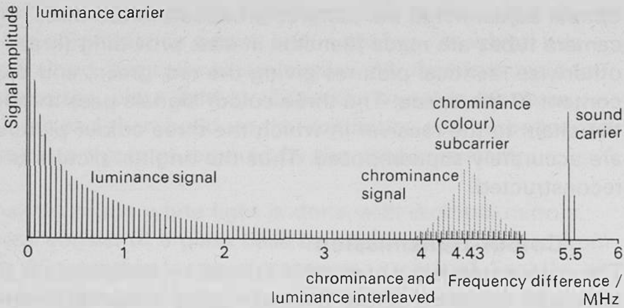


Figure 31
Interleaving of luminance and chrominance signals.

This means that the colour sub-carrier frequency must have a particular specific value at which it must be maintained with a high order of accuracy – as also must the line frequency. By maintaining this spacing to a high degree of precision a considerable improvement in overall performance is obtained. This principle of putting the chrominance modulation on a sub-carrier accurately spaced relative to the luminance carrier is an essential part of all colour television systems. The optimum spacing of the colour sub-carrier varies slightly depending on the system – for example, that for the NTSC system is not quite the same as that for the PAL system – but it is always close to an agreed value of 4.43 MHz.

Having arranged the colour sub-carrier in such a position that it causes the minimum of interference with the information on the main carrier, it is then necessary to determine the minimum amount of information to be sent on the colour sub-carrier to obtain adequate but not excessive colour performance. Many tests have been made on the acuity of human vision in colour. As a result, it has been ascertained that it is necessary to send only about one third of the detail in colour that we send in black-and-white, because the eye sees detail very sharply in black-and-white, but is much less conscious of detail in colour. We see this when matching colours with a thread of cotton. It is virtually impossible to tell whether colours match accurately by looking at a single thread; several threads or, preferably, the whole surface of a reel must be used. In 625 line black-and-white television about 5 MHz of information is sent on the black-and-white signal; in colour less than half the total bandwidth is sent, and this is divided between the two colour signals being transmitted.

The NTSC system

The NTSC, PAL, and SECAM systems differ chiefly in the transmission of the chrominance signal. I can explain this by first outlining the NTSC solution, which was the first national system to be adopted. In this case the chrominance signal is sent in two parts, the 'I' or in-phase signal and the 'Q' or quadrature signal. These are really complex signals but it is an acceptable simplification to consider them as green and red. The colour sub-carrier is divided into two components at right angles to each other. One of these components is modulated with the I signal and the other with the Q, after which they are combined to give a signal which is correspondingly modulated in amplitude and in phase.

Simple amplitude detection in the receiver would, of course, give only a composite signal containing both I and Q information. If, however, synchronous detection is used it becomes possible to resolve the phase of the signal and determine the values both in phase and in amplitude and hence determine the relative amounts of I and Q signal. To achieve synchronous detection a reference signal of the correct frequency and phase is necessary. This is provided by eight cycles of reference signal at the colour sub-carrier frequency, sent in the line suppression period, which serve to lock the oscillator in the receiver to give synchronous detection. A complete NTSC signal therefore consists of: eight cycles of reference sub-carrier sent in the line suppression period and called the 'colour burst'; the luminance signal, exactly as in black-and-white television; and the colour sub-carrier, modulated in amplitude and phase by the two chrominance signals and spaced at a distance from the main vision carrier by a precise amount corresponding to an odd multiple of half the line frequency. When no colour information is present in the picture this signal falls to zero, and it is at low level whenever the colours are unsaturated.

The PAL system

PAL was not the next proposal chronologically but it is the closest to the NTSC system. The colour sub-carrier is quadrature modulated as before but the phase of the I signal is reversed on alternate lines at the originating point and synchronously reversed in the receiver. Several means have been proposed for determining the phase of the I signal. The identification of the phase reversal switching is carried out by a change in phase of the colour burst at the beginning of each line. This change of phase is recognized by the signal processing circuit and the signal on the corresponding line is switched accordingly. This process of double phase reversal – once in the studio and back again in the receiver – has the effect that differential phase distortion on the I signal, which can occur at a number of points in the chain, is reversed in angle between successive lines. When the I signal is averaged between two successive lines the differential phase distortion cancels out. The PAL system can be used with or without a delay line in the receiver. In the former case the averaging of the two signals is done electrically; in the latter case the two signals are displayed in successive lines on the face of the colour tube and the eye does the averaging between them. If a yellow part of the picture is turned a greenish yellow in one line, by differential phase distortion it will appear as a reddish yellow in the next; at viewing distance an eye will see the sum of these as yellow.

A receiver without a delay line loses, however, a very large part of the advantages of the PAL system and in fact it would seem that the wider tolerances which PAL allows on the transmitting and distributing side would not be permissible and it would be necessary, in fact, to operate the whole of the transmitting side and to line up the receiver virtually to the same standards as are used for NTSC. If this were not done, serious spurious signals would appear in the picture displayed. In all countries at present using PAL, all receivers manufactured are of the delay line type.

The SECAM system

The third system is SECAM. This avoids the use of synchronous detection by transmitting the I and Q signals in sequence – an I during one line and a Q during the next. In order to be able to add the I and Q signals to determine the missing colour signal, each line is passed through a delay line having time equal to one line period so that, by taking the output of the delayed line and the present line, both the I and Q signals are available simultaneously. The information in each line is thus used twice.

The SECAM system went through a number of changes, from the original system which had amplitude modulation of the chrominance signal, to the later versions which used frequency modulation and had a number of variations of the signal, all designed to improve the compatibility of the system. The system at present in use in France, the USSR, and in various other countries uses frequency modulation, a varying degree of colour sub-carrier offset between one and the following line and a high degree of frequency pre-emphasis in the transmission and de-emphasis on receiver.

Choice of system

As I noted at the outset, there are many different opinions upon the merits of the various systems. But it must be remembered that something like 95 per cent of the equipment in the studio and 95 per cent of the receiver is essentially the same whatever system is used. Real differences are confined to the coding and decoding of the chrominance signal. It is not easy to assess the competing systems, and almost impossible to get agreement from the bodies concerned that any appraisal is fair and accurate; but I will offer my own views.

The NTSC receiver has the fewest components and has the lowest intrinsic cost. It is, however, rather more difficult to set up than the PAL receiver and this offsets some of the component saving. It requires a higher degree of circuit stability on the receiver. Overall, the NTSC receiver is probably slightly cheaper than the PAL. Both systems are capable of giving high quality pictures. The NTSC system requires a higher degree of accuracy in the programme distribution and transmitting services and makes exacting demands on the performance of video tape recorders. There is, however, not too much trouble in reaching this requirement and possibly the greatest advantage of the PAL system is the reduced requirements it makes on accuracy of alignment of the receiver. The PAL system also has an advantage in that recording is a less difficult problem and on the other hand programme switching is rather more complex. The PAL receiver has one less control than an NTSC receiver.

The compatibility of the SECAM signal is worse than NTSC or PAL and the inherent picture quality is somewhat lower. At high signal levels its resistance to interference is comparable to NTSC or PAL but on weak signals it is inferior to both. It does have appreciable advantage in recording but introduces very great difficulties in programme switching and mixing by reason of the fact that the chrominance signal is frequency modulated. The SECAM receiver is probably slightly more expensive than the NTSC or the PAL receiver.

Standardization

If one system of colour could be used throughout the world it would clearly be advantageous. But the world has long been divided into areas having 50 Hz and 60 Hz power supplies and this has naturally resulted in 50-field and 60-field television areas. It would seem that this division must persist for many years in colour television.

At the moment there are three systems in operation in the 50-field areas of the world and a fourth system in the 60-field areas of the world. This is an undesirable state of affairs, particularly in the light of the facilities that are now available for worldwide transmission by satellite. The situation is however coped with by the various types of standard converter that are available and digital techniques have now made these processes much more simple. The use of different standards does not therefore prevent the worldwide exchange of programmes but it does impede them. It gives particular difficulty however to people living on the borders of areas using different standards, as their receivers must be capable of operation on more than one standard.

This will clearly lead to difficulties although it should not greatly interfere with programme exchange between countries. It is possible without excessive deterioration in picture quality to trans-code from one colour television system to another, provided both are on the same line and field frequency; the equipment needed to do this is not complicated. But the problem inside the receiver is more difficult, and viewers living near frontiers of countries using different standards may require a receiver capable of accepting both systems. The best way of doing this is now being studied in several countries.

Receivers

Receivers for all systems are basically similar. The coded colour signals are transmitted and received in the normal black-and-white way. The synchronous detector in the receiver separates out the I and Q components and combines these with the luminance signal to produce the red, green, and blue signals corresponding to those which came out of the camera. These signals must then be combined to produce the colour picture.

One method of combination is to display each signal on a separate cathode ray tube with an appropriate colour phosphor and then view all three through a dynamic dichroic mirror system transmitting the three screens in rapid succession (figure 32). Very good picture quality can be obtained in this way but the device is cumbersome and superimposition is accurate only when the viewer is sitting square-on, which makes it impracticable for home viewing.

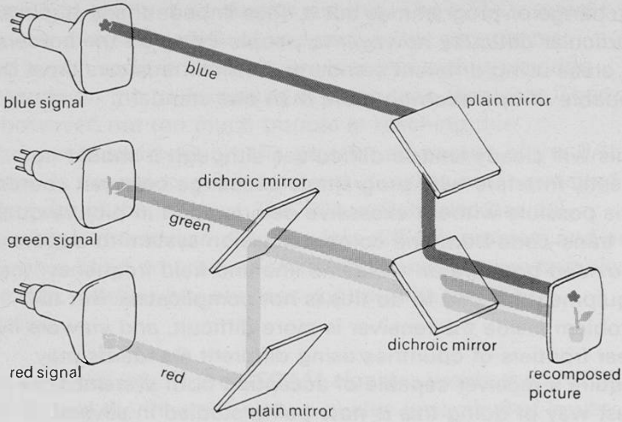


Figure 32
The colour receiver can incorporate a dichroic system of the kind used to analyse white light. The red, green, and blue images are then combined on the screen and the resulting image is seen in full colour.

Accordingly, considerable effort has been expended upon the production of cathode ray tubes able to produce a full colour picture. So far only one is in satisfactory use but it is being produced at the rate of many tens of millions a year; this is the shadow mask tube, originated by the Radio Corporation of America (figure 33). Instead of a single electron gun it has three guns each controlled by one of the colour signals. The guns are very accurately converged to 'fire' at the end-plate of the tube through a perforated plate known as the shadow mask. The latter has about 300 000 holes in it and the geometry of the tube is such that the electron beam from any one gun scans in turn through all the holes in the shadow mask in succession, line by line, but falls only on specific points on the tube face. Instead of being covered uniformly with a powder to glow white when bombarded, the tube face has 300 000 sets of three dots. In each group one dot has a phosphor which glows red when bombarded, one a phosphor which glows blue, and the third a phosphor which glows green. When the red gun is energized its electrons pass through the successive holes in the shadow mask and fall on only the red phosphor dots to build up a red picture on the face of the tube. The same applies to the green and blue. This builds up on the face of the tube three superimposed pictures which are misaligned to the extent of the spacing of the three dots in each triplet. This is a very small error and, as far as the eye is concerned, the final picture appears to be perfect.

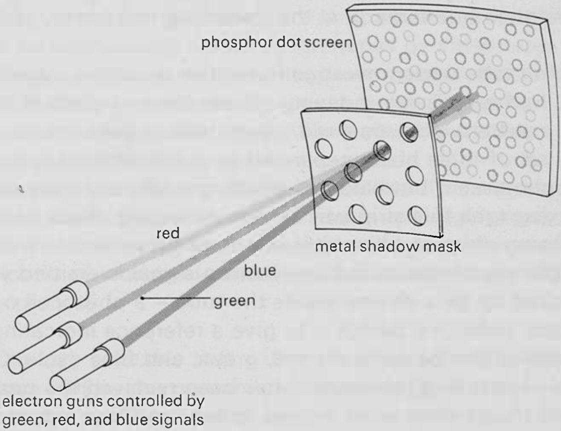


Figure 33
Shadow mask tube.

The shadow mask tube is rather expensive to make because of the accuracy with which it must be constructed. It also restricts the picture brightness, since more than four fifths of the electrons never reach the tube face but are arrested by the metal mask. There is a definite limitation to the beam current and voltage if X-ray emission is not to be excessive; too great a beam current would also heat and distort the mask. Therefore, although the shadow mask tube gives a generally adequate picture, research is in hand to find a simpler tube giving a brighter picture.

One possibility is the 'chromatron' tube in which the shadow mask is replaced by a grid and the sets of dots by vertical phosphor strips arranged in red, green, and blue sequence. By changing the polarity of the grid potential the beam can be deflected to the red, green, or blue phosphors. To do this satisfactorily demands a very rapid switching system consuming an appreciable amount of power. Such power fed to a grid of the size required for the largest tubes can give rise to radiation interference at the switching frequency.

Another tube under investigation is the so-called indexing type. In the emission indexing variety the end-plate is covered with a pattern of vertical red/green/blue phosphor bars, each set of three being separated by a thin strip of non-phosphorescent but electron-emitting material. In the photo indexing tube this strip can be non-reflecting. Each time the sweeping electron beam strikes one of these emissive strips no light signal results but an electron signal is emitted which is picked up by a device inside the tube – a photocell or electron detecting device – to give a reference indicating the position of the beam in the red, green, and blue cycle. Once a pulse of indexing information has been received the beam knows that it must send the red, green, and blue signals at definite time intervals and in a definite sequence. By this means the tube can operate with a single gun and no shadow mask. Nevertheless, the indexing circuits are complex and it is difficult to achieve adequate definition.

A further possibility is to use a tube in which three guns are aligned in a planar instead of in conical formation, firing through a grid or shadow mask plate. This may give increased brilliance and reduce manufacturing costs. In practice, great difficulties are found in damping out mechanical vibrations in a fine wired structure mounted in a vacuum. Certainly the ultimate tube for colour television has not yet been found.

Quite a number of proposals are under examination for alternative forms of picture tube but it looks as though for a long time the shadow mask tube will predominate. Since the tube was first introduced high light brightness has been increased by something like ten times by the use of improved phosphors and although there is naturally a limit to this kind of progression further improvements can be expected.

Future developments

As far as the studio is concerned, costs for colour operation are only moderately greater than for black-and-white. The increase affects only technical costs which normally are only a small fraction of the total cost of a programme production. Means for exchanging colour programmes by video tape recordings, as well as by direct landline transmissions, already exist. Satisfactory methods of programme exchange by means of recorded photographic film have not yet been resolved but undoubtedly will come; already colour transmissions have been relayed by satellite. There is no reason to doubt that, within a decade, a very high proportion of all television throughout the world will take place in colour. The growth of worldwide communications satellites will speed up the process.

It is also very probable that within the next decade colour television will be the predominant medium in all countries and that this will be supplied largely by over-the-air transmissions but with increasing contributions from wire distribution circuits and particularly from playback machines in the school, workshop, and home.

Sir Francis Charles McLean was Director of Engineering of the British Broadcasting Corporation. He is also Chairman of the Telecommunication Industry Standards Committee of the B.S.I. and during World War II was Chief Engineer in the Psychological Warfare Division of SHAEF. He is at the present time Technical Director of EVR Partnership engaged in the development of electronic video recording.

J. G. Morley

Fibre-reinforced metals

Fibres much stronger than any glass fibre promise to provide reinforcement not only for plastics but also for metals. The high-strength carbon fibres recently produced in the United Kingdom may be particularly useful.

Man's early technological development has been described in terms of the Stone Age, the Bronze Age, and the Iron Age, and this classification gives some indication of the importance of structural materials on the technology of the day. It can be argued that any radical improvements in engineering must be preceded by the development of new types of structural materials having much improved properties. Such materials should be stronger, lighter, stiffer, and preferably cheaper, than existing ones.

To make any major improvement it is necessary to start again from first principles. Every engineering material has to fulfil two main requirements: it must carry a load and it must be tough instead of brittle. A piece of it can carry a load only by suffering elastic deformation which ultimately depends on the nature of forces between individual atoms. Toughness, on the other hand, can be achieved in different ways. Wood and steel are both tough but for different reasons. Most engineering materials today are metals. As with all solids the stresses which would have to be applied to pull their atoms apart are between 10 and 100 times as great as the usable practical strength. Metals start to deform plastically at a stress very much lower than the theoretical ultimate strength, but it is this mechanism which makes them tough and thus practicable engineering materials.

Flaws in metals

All real solids must contain imperfections such as scratches produced by surface abrasion. Such flaws can have a damaging effect out of all proportion to their size and depending only on the geometrical form of the flaw. A sharp notch or crack has the ability to magnify an applied load and produce local stresses at the root of the crack many times greater than the stresses produced on the material as a whole. The stresses will be approximately proportional to the ratio of the depth of the crack to the radius of curvature of its extremity; if it is 'sharp' enough even a sub-microscopic flaw can have a catastrophic effect.

In the case of a perfectly elastic solid such a crack can spread quite easily. Following the propagation of a crack elastic strain energy is released by the relaxation of the solid, but energy is simultaneously absorbed by the creation of new surfaces. The crack will spread if the energy released is greater than the energy absorbed.

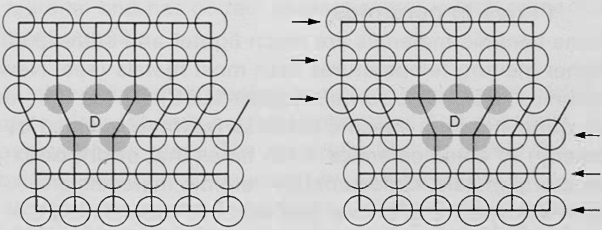


Figure 34
Dislocations spread easily through metals and are caused by incomplete rows of atoms in the crystal. When stress is applied the dislocation (D) jumps to a different position as interatomic bonds are broken and remade. Dislocations can spread in this way right through bulk metal and for this reason metals are ductile and can 'flow' in any direction. Ceramics have stronger and highly directional bonds and seldom permit the spread of dislocations except, perhaps, in severely limited directions along clearly defined planes.

In the case of a metal the highly stressed material at the crack tip starts to deform plastically thus absorbing a great deal of energy. As a result cracks propagate only with great difficulty in metals and this is why these materials are tough compared with a perfectly elastic, but brittle, solid like glass. The advantage of a metal is that is is necessary to turn only very small regions round the tips of cracks into a plastic substance in order to make this material tough. Just the right amount of the plastic component is produced in just the right place at the right time. All the rest of the material can act in an elastic manner supporting the applied load.

Nevertheless, metals have their disadvantages. They can be made harder, so that they withstand a greater load before deforming plastically, but only at the expense of making them more brittle and so there is an obvious limit to progress in this direction. Metals also become softer as they are heated so that it is difficult to make them behave satisfactorily over a wide range of temperatures. The mechanism responsible for ductility also gives rise to fatique effects. Another limitation is that the stiffness of metals and their densities tend to be linked; a stiffer metal is proportionately heavier.

Flaws in ceramics

Some ceramic materials are much lighter and stiffer and have higher melting temperatures than most metals (see 'New ceramics', *Science Journal*, August 1966). These differences are quite startling; specific stiffness (stiffness divided by density) of some ceramics is ten times that of orthodox structural metals. Unfortunately ceramic materials are generally weak and brittle, and ways of overcoming this problem have to be found if these materials are to be of practical engineering use. I described earlier how the very high theoretical strength of a solid is reduced by the presence of flaws. With ceramics a flaw cannot be made harmless by ductile deformation as it can in metals, but ceramics without flaws should be very strong indeed. High-strength ceramics have, in fact, long been known. In 1887 C. V. Boys discovered a way of making strong fibres from fused quartz: he attached one end of a molten globule of quartz to an arrow fired from a crossbow. The resulting thin fibres turned out to be many times stronger than orthodox materials.

In 1920 Dr A. A. Griffith produced a classic paper on the strength of solids. He suggested that the weakness of glass and other brittle solids was caused by the presence of naturally occurring sub-microscopic flaws, and showed that these flaws could propagate and cause failure of the material under quite low applied stresses. This theory accounted for the high average strength of thin fibres containing a small quantity of glass since there would be a much lower probability of a serious flaw being present in such a small piece of material.

Using high-strength ceramics

Of critical importance in the technological application of high-strength ceramics is the knowledge of whether the inevitable flaws are randomly distributed throughout the body of the material or occur only on the surface. Surface flaws are likely to be much easier to control and eliminate than flaws occurring naturally within the bulk material. One way to examine the situation would be to produce an experimental sample of a brittle material without a surface. Glass is a good example of a brittle material and has certain characteristics which make this experiment possible.

At high temperatures all glasses transform into liquids with a viscosity which depends on the temperature and chemical composition of the glass. The important feature is that the composition, and hence the viscosity, can be varied continuously over a very wide range. It is possible to produce a composite rod consisting of a central core of one glass surrounded by a coating of another (see 'Fibre optics', *Science Journal*, October 1965). The coating material can be such that, when heated to a relatively low temperature, it transforms to a viscous liquid whilst the core remains a brittle solid. Since there will be some diffusion at the interface between the two compositions there is no point at which the solid can be said to end and the liquid begin. Very high tensile strengths have been measured with experimental samples produced in this way. These and other experiments have shown the critical importance of the condition of the surface on the strength of ceramic materials.

It has been found that most flaws are produced by surface mechanical damage, but other types of surface flaw have also been found. These observations seem to apply to all ceramic materials and it is feasible to make quite large pieces of ceramic very strong indeed merely by removing these surface flaws. A photograph has been taken showing a crystal of sapphire (aluminium oxide) being bent elastically to a stress of about 8×10^9 N m^{-2}. Its high strength – many times stronger than the strongest steels – is due solely to its very smooth, flaw-free surface.

If these arguments about the effects of flaws apply to other ceramics, and there is every reason to believe that they do, this opens the possibility of producing a great variety of ceramic materials many times lighter, stiffer, and stronger than present day metals. But there remains the need to make the bulk material tough instead of brittle.

Although ceramics can be produced in this very high strength form it is at present hopeless to consider using such materials in bulk for general engineering purposes. Their surfaces would be too easily damaged, resulting in a catastrophic loss of strength. They would be quite impractical from the point of view of impact strength, and problems of joining and general handling damage could not be solved by any currently feasible means. Even the abrasion caused by grit and dust particles in the atmosphere could not be tolerated. Accordingly the engineering use of high-strength ceramics at first sight appears to pose insuperable technological problems but, in fact, these were solved many years ago in the manufacture of glass-fibre-reinforced plastics.

The latter materials posed similar problems. Glass fibres, although initially very strong, are easily damaged by mechanical contact; but, if they are stuck together with a synthetic resin, the fairly soft resin prevents them from damaging each other and also protects them from damage from outside. In the same way ceramic materials can be prepared in the form of high-strength fibres which, when stuck together, remain strong and also support the stresses which may be applied to the composite block in which they are embedded.

Toughness

The properties of structural materials, such as weight, melting point, and elastic modulus, depend only on the type of atoms forming the solid and the forces between them; but toughness, which is essentially a process for preventing crack propagation, can be brought about by a variety ot means. As I have described, if the tip of a crack can be blunted its tendency to propagate through the body of the material will be drastically reduced. In the case of a metal this is achieved by plastic deformation which starts at the crack tip where the stresses are highest and absorbs energy thus hindering propagation of the crack (figure 35). Other mechanisms can lead to the same result.

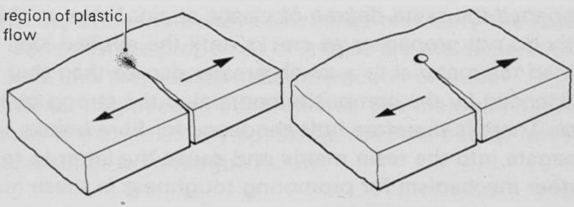


Figure 35
Crack propagation in a metal is reduced by plastic flow, and in the case of high tensile, brittle metals can be arrested by drilling a hole at the tip of the crack.

One example is to allow considerable elastic deformation, thus reducing the curvature and hence the stress-raising ability of the crack tip. This happens in the case of rubber, a very tough material. Another way is to let a crack propagate until some element in the structure is fractured. It can be arranged that at this stage the conditions are no longer suitable for crack propagation. The best known example of this is the crack-stopping hole in a metal plate. The sharp crack can propagate but, when it reaches the predrilled hole, the radius of curvature of its tip effectively becomes that of the hole and its progress is stopped.

A fibre-reinforced matrix is another example of a material in which cracks are blunted on a macroscopic scale. In such a system the fibres are subjected mainly to tensile stresses directed along the axis of the fibres. These stresses are generated by quite modest shear stresses acting at the fibre/matrix interface. High point loadings on the fibre surfaces are avoided, so that strong but brittle fibres can remain undamaged and protected by the matrix. If one fibre does contain a flaw, that fibre will break as the load is applied but, providing the matrix and interface have the appropriate characteristics, the fracture will not propagate further and damage adjacent fibres or cause them to fail in turn.

In glass-fibre-reinforced plastics the fibres are much stiffer than the resin matrix, but under load both components experience the same degree of elastic strain. Flaws in the matrix do not propagate as cracks until the applied load has strained the material to a much greater degree than that experienced by the composite containing the strong glass fibres. There is therefore little tendency for fibre breaks to propagate into the resin matrix and cause the latter to fail. Another mechanism for promoting toughness in these materials is generated by the weak fibre/matrix bond which can cause cracks to be deflected and rendered harmless.

Fibre-reinforced plastics are realistic structural materials with many attractive properties. They are in widespread use and an extensive technology has been built up around them. They can be fabricated very easily. This has enabled quite large structures such as boat hulls to be manufactured in a relatively simple way, and the advantages of low temperature, low pressure fabrication can also be applied to small precision components. Rolls-Royce has used glass fibre in major portions of lightweight lift jet engines with thrust-to-weight ratios of 16:1 or more. Other things being equal, lighter, stronger, stiffer, and cheaper fibres will lead to even more attractive fibre-reinforced plastics. However these materials still have their limitations. They cannot be used at high temperatures and the soft matrix of plastics needs to be protected from erosion under adverse conditions.

Summarizing the present position, there are two alternative routes to tough engineering materials. In the first group crack stopping is brought about by changes in material properties taking place on a very small scale around the tips of stress-raising cracks or flaws. The only practical examples of this group are metals. In the second group cracks are initially allowed to grow but are then prevented from propagating further by the material's macroscopic structure. Practical examples of this group are glass-fibre-reinforced plastics and biological materials such as bamboo.

Both types of material have their advantages and disadvantages. Metals are very efficient in that the crack-stopping component is generated in microscopic quantities just when and where it is needed, while the rest of the material continues to behave elastically and carry the applied load. Generally the properties of metals are the same in all directions and this assists the construction of three-dimensional engineering components. The major limitation of metals is that, relatively speaking, they are heavy, not very stiff, and suffer from fatigue failure.

The great advantage of a fibre-reinforced system is that the fibres can be chosen purely on properties such as low density and high stiffness; their brittleness can be ignored. But these materials are inefficient in that the second, crack-stopping, component (the matrix) has to surround all the fibres, extend throughout the bulk of the material, and of course be present all the time. A further limitation of such materials is their anisotropic nature. This is no handicap when good properties are needed only in one direction, as in the case of a fishing rod. But when isotropic sheets are required the properties are reduced to about one third of the unidirectional strength and stiffness. A three-dimensional material of this type, able to bear loads equally in all directions, is not likely to be a practicable possibility.

Whisker crystals formed from low density ceramic materials have for some years been under consideration as possible reinforcing fibres. They are very small, single crystals usually

A fibre-reinforced matrix is another example of a material in which cracks are blunted on a macroscopic scale. In such a system the fibres are subjected mainly to tensile stresses directed along the axis of the fibres. These stresses are generated by quite modest shear stresses acting at the fibre/matrix interface. High point loadings on the fibre surfaces are avoided, so that strong but brittle fibres can remain undamaged and protected by the matrix. If one fibre does contain a flaw, that fibre will break as the load is applied but, providing the matrix and interface have the appropriate characteristics, the fracture will not propagate further and damage adjacent fibres or cause them to fail in turn.

In glass-fibre-reinforced plastics the fibres are much stiffer than the resin matrix, but under load both components experience the same degree of elastic strain. Flaws in the matrix do not propagate as cracks until the applied load has strained the material to a much greater degree than that experienced by the composite containing the strong glass fibres. There is therefore little tendency for fibre breaks to propagate into the resin matrix and cause the latter to fail. Another mechanism for promoting toughness in these materials is generated by the weak fibre/matrix bond which can cause cracks to be deflected and rendered harmless.

Fibre-reinforced plastics are realistic structural materials with many attractive properties. They are in widespread use and an extensive technology has been built up around them. They can be fabricated very easily. This has enabled quite large structures such as boat hulls to be manufactured in a relatively simple way, and the advantages of low temperature, low pressure fabrication can also be applied to small precision components. Rolls-Royce has used glass fibre in major portions of lightweight lift jet engines with thrust-to-weight ratios of 16:1 or more. Other things being equal, lighter, stronger, stiffer, and cheaper fibres will lead to even more attractive fibre-reinforced plastics. However these materials still have their limitations. They cannot be used at high temperatures and the soft matrix of plastics needs to be protected from erosion under adverse conditions.

Summarizing the present position, there are two alternative routes to tough engineering materials. In the first group crack stopping is brought about by changes in material properties taking place on a very small scale around the tips of stress-raising cracks or flaws. The only practical examples of this group are metals. In the second group cracks are initially allowed to grow but are then prevented from propagating further by the material's macroscopic structure. Practical examples of this group are glass-fibre-reinforced plastics and biological materials such as bamboo.

Both types of material have their advantages and disadvantages. Metals are very efficient in that the crack-stopping component is generated in microscopic quantities just when and where it is needed, while the rest of the material continues to behave elastically and carry the applied load. Generally the properties of metals are the same in all directions and this assists the construction of three-dimensional engineering components. The major limitation of metals is that, relatively speaking, they are heavy, not very stiff, and suffer from fatigue failure.

The great advantage of a fibre-reinforced system is that the fibres can be chosen purely on properties such as low density and high stiffness; their brittleness can be ignored. But these materials are inefficient in that the second, crack-stopping, component (the matrix) has to surround all the fibres, extend throughout the bulk of the material, and of course be present all the time. A further limitation of such materials is their anisotropic nature. This is no handicap when good properties are needed only in one direction, as in the case of a fishing rod. But when isotropic sheets are required the properties are reduced to about one third of the unidirectional strength and stiffness. A three-dimensional material of this type, able to bear loads equally in all directions, is not likely to be a practicable possibility.

Whisker crystals formed from low density ceramic materials have for some years been under consideration as possible reinforcing fibres. They are very small, single crystals usually grown from a vapour phase, and the nature of the growth process tends to produce a stepped surface on the crystal. It has been shown that, if the corners of these steps are sharp enough, the steps will act as a flaw causing failure of the whisker crystal at quite low stresses. However, very small whiskers – one or two microns in diameter and a few millimetres long – tend to have very small growth steps and therefore to be very strong indeed. In the past they have been very costly to produce but more economical production processes now seem possible. Costs are still likely to be several times greater than more orthodox materials, however. Quite advanced techniques are needed to align such small components in a preferred direction. This is necessary in order to produce an adequate content of whiskers, just as for a matchbox to hold its stated contents the matches must be aligned.

Continuous fibres would obviously overcome most of these problems. High-strength, high-modulus, continuous ceramic fibres can now be produced and they are being used to manufacture fibre-reinforced synthetic resins having strength-to-weight and stiffness-to-weight ratios many times greater than those of metals and orthodox reinforced plastics.

It is tempting to try to reinforce a metal with strong, stiff, ceramic fibres and so get the best of both worlds in some sort of composite structure. The surface properties of such a composite would be those of the metal matrix, whilst the low density and high stiffness of the fibres would improve the properties of the composite compared with the unreinforced metal. Provided it is possible to make the fibres carry the major portion of an applied load, a fibre-reinforced metal will be of structural value at much higher temperatures than could be attained by the unreinforced matrix alone; and provided the high strengths of ceramic fibres can be retained, the composite material will be much stronger than conventional alloys.

It is, of course, quite feasible to reinforce a metal matrix with metal wires. Unfortunately, the properties of metals in wire form do not approach those of whisker crystals and continuous ceramic fibres. Practically nothing is gained in specific stiffness or specific strength by using metal wires instead of solid metal. These arguments apply to both 'constructed systems', where metal wires are put into a metal matrix and also to 'grown systems' in which needle-like crystals of intermetallic compounds are grown by the controlled cooling of suitable alloys. However, when the precipitating needle-like crystals are formed from ceramic materials, such as metal carbides, significant increases in specific stiffness as well as increased temperature capabilities are attainable. It is possible to align such needles in preferred directions during growth, and this is attractive from the point of view of convenient fabrication.

There is one particular application in which metal-wire-reinforced metals could be of considerable value. Many refractory metals suffer severe rates of oxidation when heated to a high temperature in air and it is difficult to protect them with an impervious coating having complete integrity under all conditions. The smallest hole will rapidly cause catastrophic failure of the whole block of material. But, if the high temperature metal were incorporated in the form of wires in a less readily oxidized matrix metal it would be protected from oxygen and from catastrophic failure following exposure of any one wire. The matrix would not have to carry an appreciable load and could be alloyed for maximum chemical resistance.

Fibre-reinforced metals differ from reinforced plastics in points of detail. With reinforced plastics there is a large difference in elastic modulus between the fibres and matrix and this causes the fibres to carry a high proportion of any applied load. In most fibre-reinforced metals this modulus disparity is much smaller but the fibres can be loaded by ductile deformation of the matrix. This mechanism could be of considerable importance in reducing 'creep' – the slow deformation of metals under an applied load at high temperatures. Creep is a major factor limiting the maximum temperature at which metal can be used.

However, there are various unknowns. Can strong ceramic fibres be incorporated in a metal matrix without degrading the strength of the fibres or introducing undesirable features into the metallurgy of the matrix? What are the general characteristics of these materials, what are their limitations, and what factors control these limitations? In order to try to answer such questions experimental fibre-reinforced metal composites have been manufactured.

Much of the basic work on this subject has been with composites of aluminium reinforced with fibres of silica (SiO_2). Such composites can be fabricated in experimental quantities and control can be exercised over a large number of the parameters which have to be considered for an understanding of the general characteristics of fibre-reinforced metals.

Tensile strength of fused silica fibres is independent of fibre thickness and largely independent of manufacturing conditions, but the fibres are easily damaged and their strength is highly dependent on the test conditions. Fibres tested at room temperature in air have average strengths in the region of 6×10^9 N m^{-2}. At temperatures in excess of about 300 °C their strength is both time and temperature dependent. However, the strengths of silica fibres are many times greater than those of any aluminium alloys and this difference is accentuated at high temperatures.

It has been found possible to manufacture aluminium/silica composites by a two-stage process. The fibre itself is produced by drawing it off as a fine filament from a rod of silica passed through a hot flame. The filament departs at high velocity and passes through a small reservoir of molten aluminium which solidifies as a coating with a typical thickness of 0.025 mm. The soft aluminium protects the silica from damage or contamination. Coated fibre is wound on to a drum and temporarily fixed with a mild adhesive. Finally, sections are cut from the drum, flattened to the desired shape, cleansed of adhesive and aluminium oxide, and finally hot-pressed until the aluminium coatings bond to form a continuous matrix. Distances between fibres can be varied by controlling the fibre diameter and the thickness of the coating. The uniformity of the aluminium coating layer leads to a uniform distribution of fibres throughout the composite, and hence to optimum mechanical properties. Sheet material similar to plywood can be manufactured by arranging multiple layers of fibres.

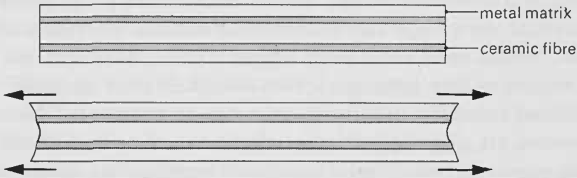


Figure 36
Fibre-reinforced metal comprises very stiff and light fibres in a matrix of metal and seeks to combine the best properties of both types of material. Presence of the fibres does not inhibit local plastic flow of the metal matrix and thus allows small scale deformation at the tips of cracks. On the other hand, the reinforcement prevents any large scale flow or creep.

Strength of this experimental material, measured in the direction of fibre alignment, has proved to be much greater than that of conventional aluminium alloys. As is the case with fibres alone, this difference is increased at high temperatures. Moreover, recent experiments have demonstrated that this experimental material has an impact strength comparable with aluminium alloys even though half its bulk consists of brittle ceramic fibres. Again, its mode of fatigue failure is interesting. It has been found that fatigue cracks in the metal matrix do not propagate and damage fibres but are deflected by them. This increases the fatigue strength of the composite by a considerable margin over that of the matrix alone. If it proves possible to combine orthodox improvements to the fatigue strength of a metal with the extra benefits which can be gained from fibre reinforcement then composites with very high fatigue strengths may eventually be produced.

Ceramic-fibre-reinforced metal systems

A very large number of possible ceramic-fibre-reinforced metal systems exist in addition to the silica-fibre-reinforced aluminium system. When all possible composite systems of apparent engineering importance are considered the number runs into thousands, even excluding dispersion-hardened metals such as TD (thoria dispersed) nickel. In the United States a broad programme of research by government establishments, the armed forces, and industry has achieved advances with such composites as matrices of silver and other metals reinforced with single crystal 'whiskers' of aluminium oxide (alumina), tantalum/tantalum carbide solidified from the melt in such a way that the carbide whiskers are aligned preferentially in the direction of the main applied stress, and a variety of composites containing fibres of beryllium, carbon, silicon carbide, and boron.

Probably the greatest research effort in the United States has been devoted to boron fibres. Their characteristics can be extremely attractive, but no way has yet been found to produce them either rapidly or cheaply. The usual manufacturing process involves depositing atoms of boron from vapour on to a very hot substrate, typically a 0.012 mm diameter tungsten wire heated electrically as in a lamp bulb. The boron not only coats but also diffuses into the metal substrate. A typical boron filament has a diameter of 0.1 mm, an ultimate tensile strength of the order of 3×10^9 N m^{-2}, an elastic modulus of 40×10^9 N m^{-2}, and a continuous length of up to 600 m; one kilogramme of it has a length of some 4 km. Initial costs were in the region of \$3000 per kg but prices have now been reduced to about one fifth of this figure. Four divisions of United Aircraft Corporation have fabricated sample hardware using boron fibre: Pratt and Whitney Aircraft have metallized fibre by plasma spraying it with aluminium or titanium and then produced bulk material (having 50 per cent of the modulus of the original fibre) for engine compressor discs, shafts, stator cases, and blades; Sikorsky are seeking to apply such fibres to helicopter rotor blades, an especially promising use; United Technology Center has drawn boron fibre through a bath of epoxy resin and wound it to form rocket motor cases of oblate spheroidal shape; and Hamilton Standard have fabricated gas pressure vessels from boron fibre wrapped round a bladder of 0.012 mm thick nickel.

In Britain most of the emphasis has been on carbon fibres. Carbon filaments have been available for many years – they formed the filaments of early electric lamp bulbs – but are used more for thermal insulation than for bearing structural loads and they seldom have tensile strength greater than 0.7×10^9 N m^{-2} nor a Young modulus in excess of $35–70 \times 10^9$ N m^{-2}. But the immense strength, and generally outstanding physical properties, of small graphite whiskers has emphasized the potential of carbon for composite development and indicated the need to produce long filaments with superior properties.

Large scale production of such filaments has now been achieved. The method adopted is to make a suitable man-made organic polymeric fibre and subject it to a very closely controlled heating process to a maximum of 3000 °C. The result is a continuous fibre of pure carbon in which the hexagonal crystallites of graphite are preferentially aligned along the axis of the fibre instead of having a random distribution. In consequence the material properties are outstanding. Each fibre has a diameter of the order of 0.008 mm, a specific gravity of about 1.8, tensile strength from 1 (average) to over 3×10^9 N m^{-2}, and Young modulus of 40×10^9 N m^{-2}. The low density results in the specific strength and specific modulus being particularly good. Strength at 200 °C is only about 20 per cent lower than the room temperature figure. It is also worth noting that the RAE have announced no loss in properties of their carbon fibre material after immersion for one month in water; penetration of liquids is one of the hazards to be avoided with a reinforced composite material. In a hank the fibres appear black and silky, highly flexible, and difficult to break (see *Science Journal*, July 1966, pages 3 and 14, and August page 16).

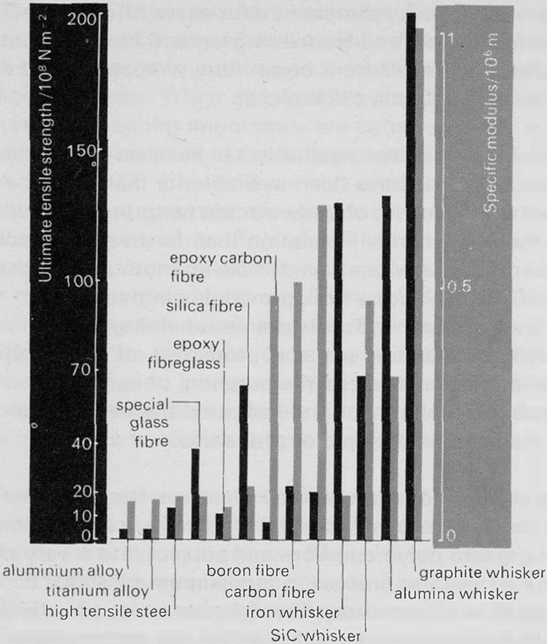


Figure 37
Tensile strength and specific modulus (ratio of stiffness to weight) of contrasting structural materials.

To use carbon or boron fibres it is necessary to build them into a composite structure and it should be emphasized that generally speaking it is unsatisfactory and uneconomic to do this by orthodox machining techniques such as are used with metals. Ideally the fabrication of the material should be by the same process as that which produces the component in which it will be used. Only in this way can the full potential of such material be realized. The most widely used raw material is known as prepreg sheet. This sheet material, typically up to a tenth of a millimetre in thickness, consists of a parallel array of fibres impregnated with a controlled quantity of polymeric material. The relative proportions of fibres and polymer, and the thickness and uniformity of the sheet, are closely controlled. A large number of sheets placed one above the other is known as a preform and such an assembly can be placed in a mould and be pressed together and heated to produce the finished component. In this method of manufacture the mould has the shape of the finished part which can be quite complex. The individual sheets of preform are cut into appropriate shapes to fit the mould, and the orientation of the fibres within each sheet can be controlled to give the required strength and stiffness in the finished part. Manufacturing techniques of this type have been used by Rolls-Royce to produce glass-fibre-reinforced compressor blades for the RB.162 lightweight lift jet engine.

Rolls-Royce have given the name Hyfil to the first series of composite materials incorporating new types of fibres. Various resins have been used to provide the bond but in all cases the proportion of fibre exceeds 40 per cent by volume and 50 per cent by weight, and the fibre properties are substantially retained. These materials are similar to the laminated glass fibre in epoxy resin used extensively in the RB.162 lift jet but the fibre itself is both stronger and stiffer.

The mechanical properties of the first grades of Hyfil are encouraging enough for us to be optimistic regarding the eventual application of these materials. To a first approximation the achieved specific strength and modulus are five times as high as typical figures for titanium and four times the best previously achieved with other reinforced plastics. However the relatively soft polymeric matrix requires protection against erosion from sand and other foreign particles. When these materials have reached a suitable stage of development they are expected to be used in the cooler parts of large gas turbines.

The compressor of a gas turbine has to produce a large volume of air at high pressure. It must be as light as possible for the work it has to do and a very high proportion of the mechanical energy applied to the rotor must be transformed into useful energy of the gas stream. One way towards improved aerodynamic design of compressors and fan stages makes use of longer, thinner, and narrower blades. Such blades are more prone to vibration and flutter problems than are more conventional, smaller blades. Although this problem can be avoided by linking blades together with struts this action reduces the aerodynamic efficiency below

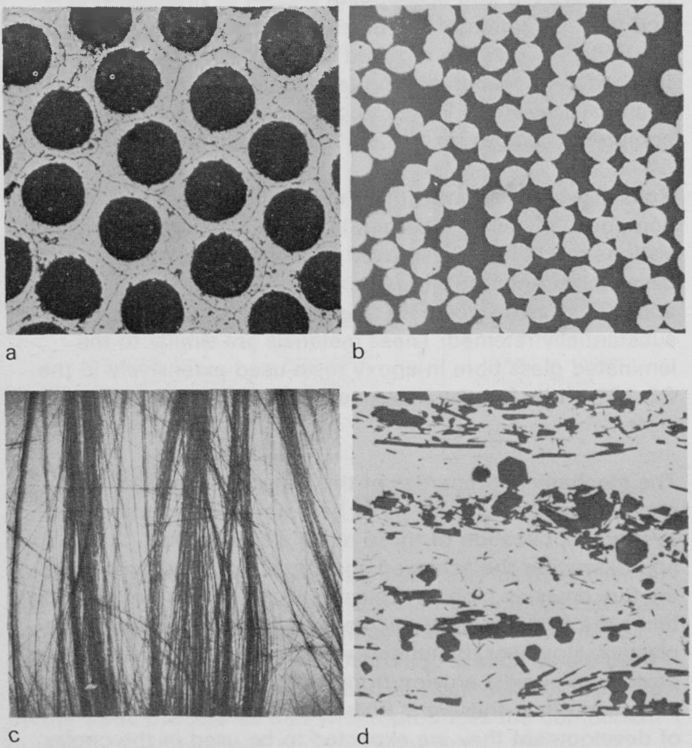


Figure 38
Electron micrographs showing:
a silica-reinforced aluminium;
b carbon fibres in epoxy resin;
c carbon fibres;
d alumina whiskers embedded in a nickel matrix.

the level that would otherwise be achieved. One way to improve the aerodynamic efficiency of a compressor, whilst at the same time avoiding flutter problems, is to manufacture its blades in a lighter and stiffer material and it is for this type of application that the new Hyfil composites offer great promise.

But we are only at the beginning of the development of this new materials technology and a great deal remains to be done.

A considerable body of experience and understanding will need to be built up before large scale applications can be contemplated. Costs are likely to remain high compared with more orthodox materials but low enough to enable the new composites to be cost effective in aircraft applications.

The quantities of these new composite materials which industry will eventually require is anybody's guess because there is no doubt that they will come to be used, and to an ever increasing degree, in all branches of advanced engineering.

Professor John Godfrey Morley is Special Professor, Wolfson Institute of Interfacial Technology, University of Nottingham. Previously he was Head of Composite Materials Research, Rolls-Royce Advanced Research Laboratory, Derby, and before that Head, Physical Properties Group, Pilkington Brothers Research Laboratory.

Acknowledgements

Figure 38: (a) Rolls-Royce (1971) Ltd.; (b) and (c) Procurement Executive, Ministry of Defence, RAE, Farnborough; (d) Dr C. A. Calow, AWRE, Aldermaston.

Gordon Newstead

The homopolar generator

With a rotor weighing 40 tonnes and an output of 1.6 million amperes, the homopolar generator in Australia is the largest electrical generating machine in the world. It is now being used to generate very intense magnetic fields. Homopolar generators may eventually find a range of industrial and domestic applications.

At the Australian National University, Canberra, is probably the largest and certainly the most unusual electric generating machine in the world. The successful construction of this machine is a landmark in more than a century's investigation and controversy about the practicability of these machines which are known as homopolar generators. Essentially they consist of two enormous rotors in the form of discs – each weighing about 40 000 kilogrammes – which spin in a magnetic field. A voltage applied across the discs – between their centres and their rims – causes them to rotate for the same reasons as the rotor of a conventional electric motor rotates in its magnetic field. The homopolar generator is not, however, used as a motor. When the discs reach a speed of some 900 revolutions per minute – with the periphery moving at half the speed of sound – they are stopped rapidly which, in practice, means about one second. The result is a surge of electrical current reaching a peak value of more than 1.5 million amperes. The energy transfer involved in this process is roughly equivalent to bringing a Boeing 707 travelling at about 240 kilometres per hour to rest in one second.

The generator is brought up to speed by energy supplied from the electric mains, a process which takes about ten minutes. During this time electrical energy is converted into the kinetic energy of the spinning rotors. The function of the generator is thus primarily to store electrical energy and release it in a short burst as a huge pulse of current when the generator is heavily loaded. An electrical capacitor bank capable of storing 10 megajoules of electrical energy is currently considered very large; the homopolar generator is capable of storing 500 megajoules. If this energy could be completely converted into mechanical energy it would be capable of raising a 1000 kg weight to a height of about 50 kilometres.

This, of course, is not the purpose of the homopolar generator. Large bursts of electrical energy are required in many fields of physics, particularly for the generation of very high magnetic fields. This now seems likely to be the most important application of pulsed homopolar generators although continuous homopolar generators may eventually find a wide range of application in industry and consumer products.

Development of homopolar generators

Michael Faraday invented the homopolar generator in October 1831 (figure 39). His device was the first electrical machine capable of producing an appreciable electric current. Methods of generating electricity until then had used either frictional effects to produce 'static' electricity (for which even now there are few applications) or chemical energy in an electric battery.

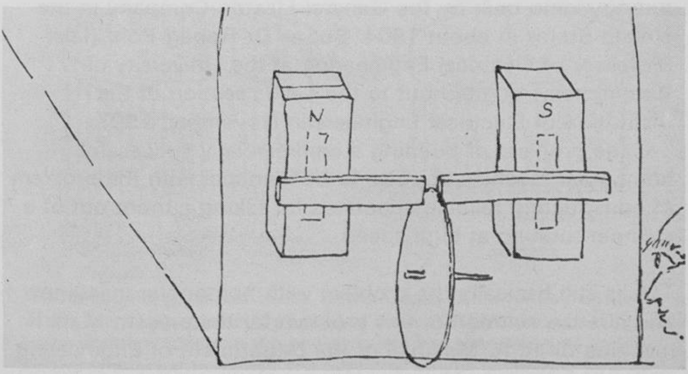


Figure 39
The first homopolar generator. Taken from Faraday's diary of October 1831.

Faraday's invention is important because it was the first method proposed for converting mechanical energy into electrical energy and his discovery has formed the basis of modern electrotechnology. However, the major result which followed from Faraday's invention was not the development of homopolar generators and motors but the discovery of a principle of electromagnetism: if a conductor rotates in a magnetic field the voltage induced in it is proportional to the rate of cutting of the magnetic flux. This 'flux-cutting rule' has been used to develop the generators, transformers, and motors now widely used. It is a curious fact of the history of science that the flux-cutting rule which was initially developed from experiments with a homopolar generator was considered by many people to be inadequate to explain its operation, and a long and often acrimonious discussion of the matter will be found in the scientific and technical literature. However, the existence of this literature brought the homopolar generator under the notice of all serious students of electromagnetism, prevented it from becoming lost and forgotten, and raised the question of its possible construction and usefulness.

A few machines were actually built. Probably the most successful of the early models was a 300 kilowatt (kW) turbodynamo built by the General Electric Company in the United States in about 1904. But as Dr Robert Pohl (later Professor of Electrical Engineering at the University of Birmingham) pointed out to the Leeds section of the Institution of Electrical Engineers in November, 1907, '. . . the problem of building a commercially successful homopolar machine appears to be identical with the problem of constructing reliable apparatus for taking current out of a cylinder rotating at high speed. . . .'

This is still basically the problem with homopolar machines and it is the solution of this problem for the present Mark II machine by R. A. Marshall of the Department of Engineering Physics, Australian National University, that has led to its current success. An earlier version of the machine (the Mark I) had liquid metal brushes and servo bearings. While this was successfully tested, it was altered to the present machine for engineering and operational reasons. In this article I shall describe mainly the Mark II machine which has been operational for over a year and is still being improved.

The Canberra homopolar generator, which is unique in many aspects, was conceived and constructed by Sir Mark Oliphant and his colleagues J. W. Blamey, P. O. Carden, L. U. Hibbard, E. K. Inall, R. A. Marshall, and B. Shenton, all of whom have worked continuously on the project for many years. I have been in charge of the project for the past two years. As we shall see, the design of a homopolar generator of the size of the one at Canberra involves working at the limits of technological ability in a variety of fields.

Basic principle

In principle the homopolar generator consists of a conducting disc or rotor spinning about an axis in a magnetic field parallel to this axis (figure 40). The output voltage is available at the terminals of the machine from two sliding

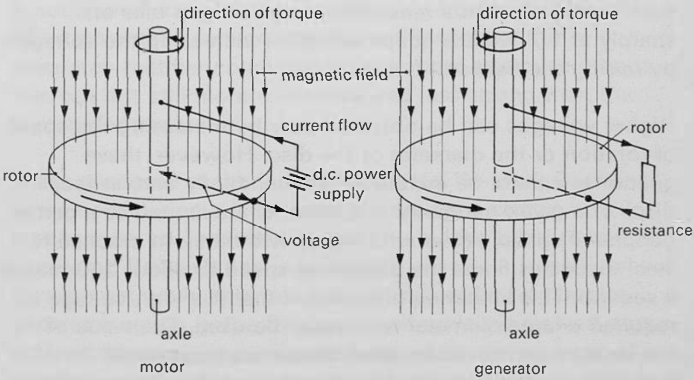


Figure 40
The homopolar generator. A d.c. power supply is fed to the disc (left) which eventually rotates at a speed of some 900 revolutions a minute. In the righthand diagram the power source has been removed and the system connected to a load. The rotational energy is released as a large pulse of current. This current produces a torque which opposes the rotation of the disc, and the disc is brought to a halt in about one second.

brushes, the inner one of which can, in theory, be a collar or part of the bearing which supports the axle.

According to the flux-cutting rule a voltage is generated between the centre of the disc and its perimeter because as the disc rotates the line of contact cuts across the magnetic field. The magnitude of this voltage increases with magnetic field strength, speed of rotation of the disc, and its diameter. In order to obtain a large voltage it is necessary to produce a large magnetic field, a high angular speed, and a large diameter disc.

In a large machine the magnetic field has to be produced by an electromagnet which itself requires electric power to drive it. There is obviously no point in making an electromagnet which consumes all, or even an appreciable fraction, of the generator's output. We therefore had to consider field magnets that consume as little power as possible and this led us to iron-cored magnets. However, because iron saturates at about 2.1 tesla (T) the efficiency of such a magnet with a reasonably large air gap falls off sharply at 1.7 T. Little scope is left for increasing the voltage by raising the field still further.

Higher voltages can be obtained only by increasing the speed of rotation or the diameter of the disc. However, these quantities cannot be increased independently and, in fact, discs of a given geometry and material are limited to a certain peripheral speed above which they will burst. In practice for steel discs this limits the peripheral speed to about 200 metres a second. This limitation also means that if a high voltage is required a large diameter rotor must be used. The value of the voltage generated by steel discs thus turns out to be 85 times the diameter of the disc in metres. A disc of two metres in diameter rotating so that its peripheral speed is 200 m s^{-1} will therefore generate 170 volts. To improve this performance we need material for the rotor which is stronger and material for the field magnet which saturates less easily. Such materials are available but the improvement resulting from their use is not great and the extra cost is considerable.

Advantages

The main advantage of the homopolar generator over other machines is that it has no windings. This makes it much stronger than a conventional machine and enables it to be used more effectively in a pulsed manner. In this method of operation, which was developed by P. Kapitza in the early 1930s at the Cavendish Laboratory, Cambridge, for producing intense magnetic fields, an alternator is taken up to speed and then heavily loaded. During the short circuit the machine develops many times its rated output as it slows down or stops. The energy it delivers is supplied from the kinetic energy of rotation. These Kapitza or pulsed type alternators are now fairly common, but the ultimate limit of what can be produced from an alternator of a given size is governed by the fact that during the current pulse the windings tend to be torn off the machine. Furthermore, alternators produce alternating current whereas direct current is usually required for the kind of experiments which need very large pulses of electrical current. Direct current generators use a commutator for turning alternating into direct current but this is not a satisfactory device for very large currents. Recent developments in semiconductor rectifiers make the alternator–rectifier combination look more attractive than it was but it is still more expensive and less strong than the homopolar generator.

These are some of the factors which led Sir Mark Oliphant, when considering a large energy store for a proton synchrotron, to consider the use of a pulsed homopolar generator. Basically the idea is quite simple. A large flywheel stores energy which is taken out in pulses by homopolar generator action in the flywheel itself which is slowed down or even reversed (for an inductive load) in the process. This arrangement has several attractive features. Firstly, a flywheel rotating at speeds where it is stressed up to allowable limits is one of the most efficient methods of storing energy; secondly, the low internal resistance and inductance of the homopolar generator allows the energy to be taken out in times of a second or less; and, finally, energy does not have to be transferred by shafts and keyways which in a conventional machine are usually highly stressed. Energy can thus be stored over a period of time by running the machine up to speed as a homopolar motor and can then be quickly discharged using the machine as a homopolar generator. In fact, the machine is electrically equivalent to a capacitor and is a highly economical way of achieving a very large effective capacitance.

Problems

With a peripheral speed of 200 m s^{-1} there is a considerable amount of frictional heating at the outside brush. There is also electrical heating caused by the passage of current across the contact potential difference which occurs between the brush and the rotor. The outer brush is also subject to rapid wear. Furthermore, if the disc is not reasonably round and running true the brush will have difficulty in following the surface of the rotor and it will tend to bounce about and only make contact intermittently with the disc. This will lead to arcing and further generation of heat as well as hammering of the brushes.

A possible solution to these problems is to use liquid metal brushes. In such a machine the current could be collected by jets of, say, mercury which are squirted on to the rim of the rotor. These brushes would have the advantages that they would cool rather than heat the rotor, and there would be no mechanical wear. However, mercury is expensive, its vapour is poisonous, and it easily becomes contaminated (producing a fall in its conducting properties). These last two factors would require a completely sealed machine from which even a very small leak of mercury vapour over a period of time would be an insidious danger to the health of those working near it. For these reasons mercury was not used in the Canberra machine which originally incorporated brushes of a sodium–potassium alloy (NaK). This is a liquid conductor at room temperature but it ignites in air and has to be used in an atmosphere of nitrogen. The machine was successfully completed with the NaK brushes and was known as the Mark I machine. However, after an explosion during an operation involving the handling of NaK away from the machine, it was decided that the alloy was too dangerous and required too elaborate safety precautions to make a successful operational machine. Marshall then investigated the use of carbon brushes and showed that they were feasible in a pulsed machine.

A rotor supported on an axle as shown in figure 40 requires bearings for it to run in and these present some unusual difficulties. First, they have to be accommodated within the magnetic field. This requires a hole in the structure which has to be kept small to minimize the distortion of the field. Great care is needed in the manufacture and insulation of the bearings so that the voltages induced in them do not cause currents to circulate which overheat them and to ensure that, in fact, the brush problems do not occur again in the bearings.

A steel disc orientated at right angles to a magnetic field is unstable and unless restrained will turn itself parallel to the field. Steel discs are used to reduce the path length in air of the magnetic lines of force and so facilitate the production of the required magnetic field. Generally, the restraint is provided by the stiffness in bending of the axle. The force tending to tilt the axis which has to be catered for in the Canberra machine is equivalent to holding a 200 000 kg weight on top of a 180 metre pole. Obviously it is difficult to cater for a torque of this size by increasing the strength of the axle. Oliphant suggested that the most effective solution was to apply stabilizing forces around an annulus whose diameter is an appreciable fraction of that of the rotor. He initiated work on the air bearing shown in figure 41 which was then carried through to a highly successful conclusion by E. K. Inall.

The restraining force is generated by air forced through small holes in the surface of the bearing so that the rotor floats on a film of air a few thousandths of a centimetre thick. There are also oil thrust and stabilizing bearings attached in effect to the axle to provide stability against vertical and lateral forces.

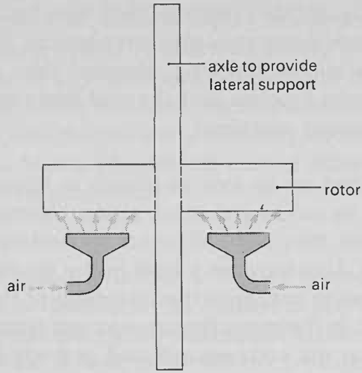


Figure 41
Air-stabilized bearing.

After the success of the air bearing Oliphant produced a new design for a homopolar generator in which all the support (lateral, vertical, and rotational) is supplied by air pads. In this design the rotor is entirely supported and restrained, like a hovercraft, on thin films of air.

The Mark II Canberra machine

The Mark II Canberra machine (figures 42 and 43) has two rotors. These rotate in opposite directions so that no net torque is transmitted to the machine foundations, although it is transmitted through the yoke of the main field magnet. The rotors are connected in series to increase the voltage. Each rotor is divided into two discs which are insulated from each other so that, if required, four discs can be connected in series to produce four times the voltage generated by one disc. The insulation is accomplished by glueing each half rotor to a rubber separating mat with epoxy resin – a courageous experiment which was carried out by Dr L. U. Hibbard, formerly of the Department. We believe that these rotors must be the largest and heaviest pieces of metal held together entirely by glue on any structure in the world.

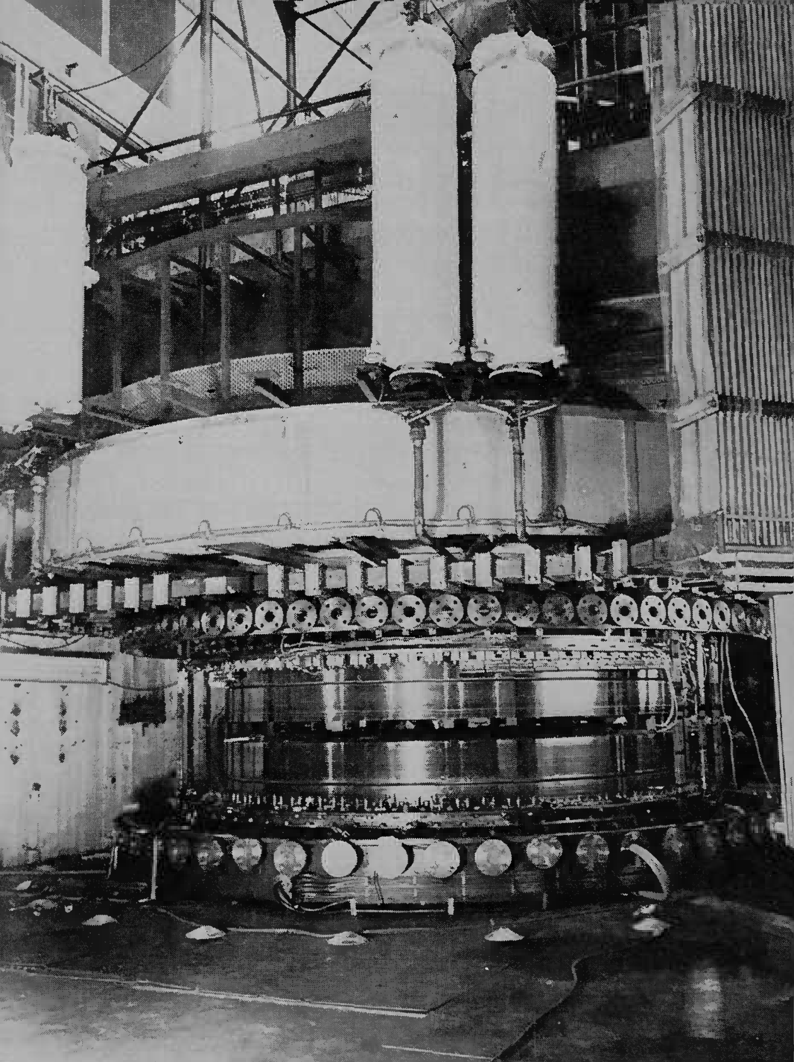


Figure 42
The Mark II Canberra machine. In this photograph the brush gear has been removed and the discs and rotors are clearly visible.

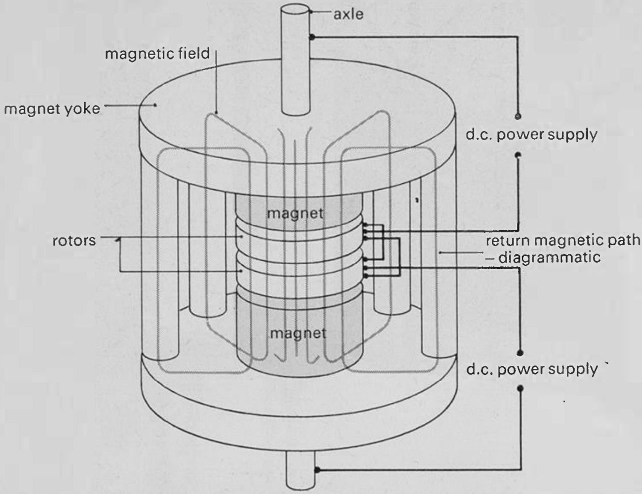


Figure 43
This schematic drawing of the Mark II Canberra shows how the rotors are connected and the arrangement of the electromagnet supported on its yoke. The magnet is located above and below the discs so that a magnetic field passes through the discs perpendicular to them.

Each rotor is 3.6 metres in diameter, weighs 40 000 kg and can rotate at speeds up to 900 revolutions per minute. At these speeds the energy stored in the rotor is 576 megajoules. In the Mark II machine most of this energy can be discharged in one second in the form of a large current pulse of 1.6 million amperes (figure 44). The machine is therefore comparable so far as forces and currents are concerned to the largest alternator now in existence – that delivering 500 megawatts. With continual improvements (mainly in operational features and instrumentation) the Mark II machine has operated now for more than a year and has produced about 600 current pulses of which 50 were more than 1.5 million amperes. The maximum energy taken from the generator to date is 250 megajoules. Smaller currents for a longer time can also be produced. For example, in one particular experimental set up, the machine has regularly been producing 25 000 ampere pulses of 30 second duration.

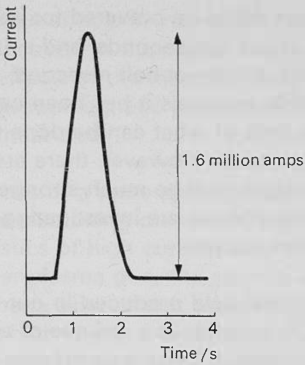


Figure 44
Large pulse of current released from the homopolar generator.

If the machine is required to operate in this way some of the brushes are removed to reduce the frictional heating. The brushes themselves are the usual type of copper–carbon sinter and are applied by compressed air during the pulse.

High magnetic fields

The Canberra machine was originally planned for use as the energy store for a large proton synchrotron accelerator. When this project was abandoned some years ago it was decided to complete the machine for use in plasma experiments. It is not now planned to use the machine immediately for this purpose although it may eventually be used for plasma work. The most fruitful application at present seems to be in the production of very high magnetic fields for use in experiments in solid state physics.

In generating strong magnetic fields a distinction must be made between fields which last for only a few milliseconds and those which last for several seconds. The former require difficult instrumentation and are not suitable for the study of phenomena which take time to reach equilibrium; the latter are for all practical purposes constant. P. O. Carden recognized that the homopolar generator was capable of powering the strongest magnets that could be made. Calculations showed that a 15 T magnet could be powered for about 30 seconds, a 30 T magnet for about five seconds, and a 100 T magnet, if it could be made, for about half a second. Because of the strength of available materials it has been calculated that 30 T is about the limit of what can be done with existing techniques and materials. However, there are indications that fibre-reinforced copper will be much stronger and still have a high conductivity and we are investigating the use of such materials in magnet design.

At present the highest field produced in our laboratory is 16.5 T for about 20 seconds in a volume of about 5.5 cm diameter × 10 cm using a Bitter type magnet manufactured for us by A. D. Little Inc. We are at present designing in conjunction with the Oxford Instrument Company a magnet described as a 30 plus tesla magnet. This magnet will produce at least 30 T for a few seconds and may, with some relaxation of the factors of safety or life of the magnet, produce as much as 40 T. This magnet, for which Martin Wood of the Oxford Instrument Company is mainly responsible, is an interesting design and consists of a small 15 T magnet inside a larger 15 T one. This type of cascading cannot go much further and, in the last analysis, may achieve little more than improved control over the current distribution in the coil. Nevertheless, it appears to be a satisfactory design and we expect to have it operating in about a year. This design is probably about the limit of what can be achieved with conventional materials and if higher long-time fields are to be produced new materials and/or techniques are required.

We intend to use these magnets to study the properties of matter in high magnetic fields. The facilities of the laboratory are available to other universities and institutions, and we propose to develop our own group working in this area. The 15 T magnet has already been used by Professor R. Street and B. Munday of Monash University to study field-induced transitions in chromium by observing changes in the velocity of sound in specimens subjected to the field. Another group at the University of Adelaide proposes to study cyclotron resonance effects in indium antimonide. At these high fields the electrons in the crystal behave like a dense plasma and the effects of the magnetic field on the propagation of electromagnetic waves through the material will be measured.

Other investigations of a more speculative nature and needing large currents are also being planned. These include a study of the characteristics of high current arcs, the construction of a very high power plasma gun, the physics and technology of brush phenomena as well as a search for a more efficient material out of which to make brushes, the use of the generator as a welding machine to resistance-weld very thick plates, and the possibility of turning the energy of the generator back into kinetic energy of a small mass which would have a correspondingly high velocity. We calculate that we have enough energy to give a mass of 2.5 kg a velocity of 10 000 m s^{-1} and we are investigating ways of doing this.

This velocity is of the order of orbital velocity and could be used for the study of missile re-entry problems. The possibility of using the generator to power a supersonic wind tunnel is also being investigated. It is also possible that a plasma experiment will be designed around the homopolar generator, but we believe that the most important contribution made so far is the demonstration that large energy stores of this type are possible and relatively cheap, and that a great deal of scientific and technological information will result from studies made with the large currents and energies available from the machine.

The future

Apart from the uses of this particular machine, what is the future for homopolar generators and motors? In answering this question it is necessary to consider the two types of machines: pulsed machines such as the Canberra homopolar generator I have described, and continuously operated generators or motors. There is no doubt that the homopolar generator is the best method of storing megajoule amounts of electrical energy when this energy is required to be released in times of the order of 0.1 to 10 seconds. Its rivals in this field are the alternator–rectifier set, the capacitor, and the storage battery. It is cheaper than all these devices and has marked operational advantages over some of them. The cost of energy storage in the Canberra machine is 0.25 cent a joule compared to 5 cents for a capacitor and 2 cents for a lead-acid accumulator. The alternator–rectifier set might have advantages of ease of control and the fact that it is a more standard device, but a machine of the same output as the Canberra machine would be at least ten times as expensive and would not be as strong.

In the continuously operating type of machine the limitations are still in the brushes. Marshall is carrying out further research into this matter along the lines of searching for better materials and trying new methods of construction. Assuming that this research is successful then there would probably be a large field of application of small homopolar motors for such things as vacuum cleaners, fans, and pumps.

Homopolar motors could be useful for driving the propellers of electric ships and, if the fuel cell becomes an economical proposition or the power-to-weight ratio of batteries increases markedly, than the long dreamed of electric motor car with homopolar motors (also used as regenerative brakes) built into the wheels, could become a reality. In all these applications the great advantage of the homopolar machine would be its robustness and simplicity.

Further development can be expected in homopolar machines themselves. For a long time people have been trying to get rid of the brushes. Because of the mis-understanding referred to earlier about how the homopolar machine actually worked there were many people in the early days who thought they had invented homopolar machines without brushes.

This led A. M. Gray to write his paper 'Impossible homopolar machines' in which he showed that machines of this type necessarily require sliding brushes. Although this was true for the technology of the day, the development and application of the phenomenon of superconductivity has rendered this conclusion false. For example, J. Volger, D. van Houwelingen, P. S. Admiraal, and J. van Suchtelen at the Philips Laboratory, Eindhoven, have developed a machine like ours except that the sliding brushes have been omitted and the disc is made of superconducting material (figure 45). A small sector of the material of the disc is made normal – non-superconducting – by subjecting it to a sufficiently strong magnetic field. This sector is made to rotate, either by rotating the magnet system which produces the field or by applying an equivalent system of alternating magnetic fields in the same way as the rotating magnetic field is produced in an induction motor. Because magnetic lines of force cannot penetrate a superconductor, all the field passes through the normal section of the disc and hence as it rotates it induces a voltage in the closed turn.

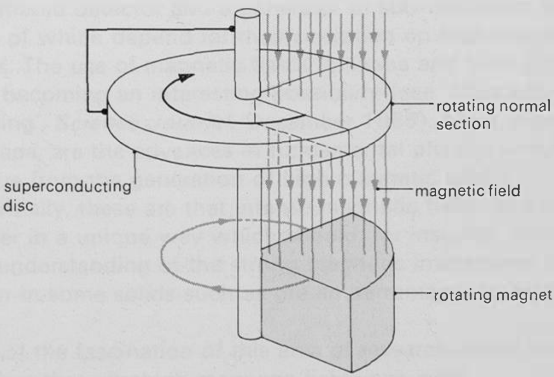


Figure 45
The brushless homopolar generator developed at Philips Research Laboratory, Eindhoven.

Homopolar generators of this sort have the advantage that they have no moving parts, but this is more than made up for by the complexity of the apparatus required to keep the disc at the very low temperatures required to produce superconductivity. Such machines are at present no more practical than Faraday's original concept but like that device they could, one day, develop into practical machines.

Emeritus Professor Gordon Newstead was formerly Head of the Department of Engineering Physics at the Australian National University, Canberra. He is currently an Associate Commissioner of the Hydro Electric Commission, Hobart, Tasmania.

D. H. Parkinson

Ultrahigh magnetic fields

Laboratory generation of yet another extreme research environment makes possible the development of devices dependent on intense fields for their operation and enables a more detailed study of the solid state to be initiated.

One of the major preoccupations of contemporary physics is a desire to create extreme environments. Over the past decade or so, physicists have been highly successful in producing extreme temperatures and pressures, both high and low, which do not occur naturally on the Earth's surface. In this article I shall discuss another of these extreme environments, the ultrahigh magnetic field.

The desire to produce very large magnetic fields is prompted partly by technical as well as by basic considerations. For example, recent advances have made it possible to develop an infrared detector and a generator of sub-millimetre waves both of which depend for their operation on high magnetic fields. The use of magnetic fields to shape and form metals is also becoming an interesting possibility (see 'Magnetic forming', *Science Journal,* December 1965). Most important, perhaps, are the advances in fundamental physics which may accrue from the generation of high magnetic fields. Essentially, these are that intense magnetic fields interact with matter in a unique way which should, for instance, increase our understanding of the strong magnetic interactions which occur in some solids such as the antiferromagnetic materials.

Part of the fascination of this area of research stems from the fact that ultrahigh magnetic fields can exert tremendous pressures and can transmit forces with the velocity of light. Their study offers the possibility of examining, with some degree of control, matter which is subjected to conditions of extreme pressure.

I shall discuss the contributions both to basic physics and to technology which this field is making in some detail after describing the basic means which are now being used to generate ultrahigh magnetic fields.

Work began on the generation of intense fields 30 years ago but only recently has this study attracted wide scale support. The most important advance in fact occurred as long ago as 1914 when H. Deslandres and A. Perot showed how a high powered water-cooled solenoid could be used to generate powerful continuous fields. The significance of their work was not appreciated at the time. However, after steady advances in several different laboratories in the intervening years, we have now seen in 1963 the setting up of the United States National Magnet Laboratory at the Massachusetts Institute of Technology (MIT). This laboratory has produced the most intense continuous magnetic fields so far attainable: steady fields well in excess of 20 tesla – about 500 000 times the strength of the Earth's magnetic field. Similar research projects are being planned in the United Kingdom, France, and the Soviet Union.

Solenoids

The basic method of producing intense fields involves passing a current of several thousand amperes through a solenoid – a coil of wire wound in uniform layers around a cylinder – usually constructed of copper. In the 'orthodox system' – which uses water-cooled copper conductors – about two megawatts of electrical power are needed to create a field of 11 T. A typical 2 MW coil will be 0.15 m long and have an outer radius of 0.15–0.2 m. Consequently, the cooling problem is severe; usually about one cubic metre of water a minute must be passed through the coil at about seven times atmospheric pressure for each megawatt of power used.

This system requires expensive equipment because the field which can be generated is proportional only to the square root of the power available. For this reason, also, a power plant which has a built-in high overload capacity – even if only for a few minutes (or seconds) is a great advantage. Consequent economic considerations rule out the use of local 'prime-movers' such as gas turbines to generate the current and it is cheaper to use the National Grid as the power source. Orthodox high power magnets usually require a d.c. supply of between 100 and 200 volts and whatever system is used to convert the grid supply to the required d.c. voltage, the cost of the power will invariably be high unless running is restricted to off-peak hours.

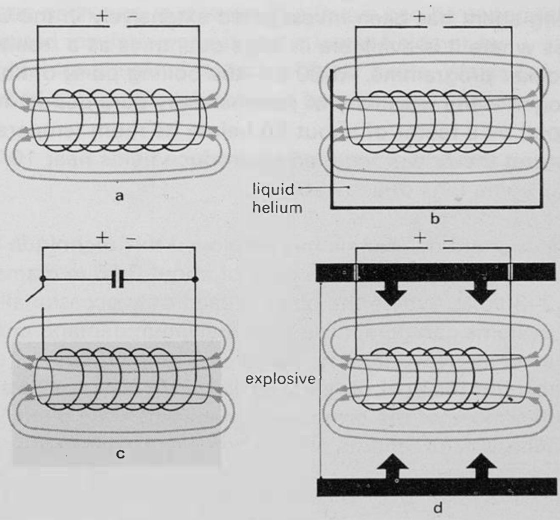


Figure 46
Four basic ways of producing ultrahigh magnetic fields. All depend essentially on passing a heavy current through a coil wound round a cylinder (a). Power consumption can be decreased by cooling the *cylinder to lower resistance or even cooling it to the superconducting* state (b). Even higher magnetic fields, though very shortlived ones, can be produced by discharging a capacitor through a coil (c). The highest magnetic fields are produced by flux compression, shown schematically in (d), in which the field is compressed by detonating explosives at the critical time.

Ways of overcoming these economic difficulties are obviously important and the first step is to cool the solenoid well below room temperature. The electrical resistance of a pure metal drops more or less proportionally to its temperature down to about 20 K where it tends to level off slowly to some 'residual resistance'. At such temperatures, therefore, a fixed amount of power will produce a greatly increased current in the coil and, hence, a greatly increased field.

The use of liquid hydrogen (as well as liquid nitrogen) as a cooling agent has been investigated extensively in the United States where it is available in large quantities as a result of the rocket programme. At 20 K – the boiling point of liquid hydrogen – the resistivity of commercially pure copper has dropped by a factor of about 50 below its room temperature value and the power required to produce fields near 10 T is only some tens of kilowatts.

H. Lacquer at Los Alamos has employed this technique to produce fields up to 10 T in coils of about 0.06 m diameter with 2–3 volts. Where the plant is used only occasionally, such systems can defeat the high maximum demand charges for electricity. Nevertheless, the overall running costs – taking into account the cost of liquid hydrogen – may not be any less than those of the orthodox system when the plant utilization is high, and its greater complexity has many disadvantages.

Problems of design in both orthodox and refrigerated coils are similar. They are principally concerned with defeating the powerful forces tending to burst the solenoid, and with extracting the heat from it. The forces are of two kinds: longitudinal ones, which are directed inwards and tend to compress the conductors and insulating material; and the more troublesome 'barrel' forces which are directed outwards.

The cooling problem is equally complex. There is an upper limit to the rate at which heat can be forced through unit surface area of a metal in contact with a turbulent cooling liquid. Above this limit a vapour film will be formed at the surface through which thermal conduction is relatively so poor that the metallic temperature will rise disastrously. This limit is around 10 MW m^{-2} and, in practice, depends very much on the construction of the cooling channels.

Two types of solenoid have been constructed – the strip-wound and the disc type. As an example of the first type, M. Wood at the Clarendon Laboratory, Oxford, designed a coil which consists of concentric layers of strip wound on its narrowest edge (figure 47). Thin insulating strip is inserted between each turn in a layer. The layers can be separated by corrugated insulating sheet at the lower fields or by individual longitudinal strips for higher fields. Water can then be forced along through the coil between the layers.

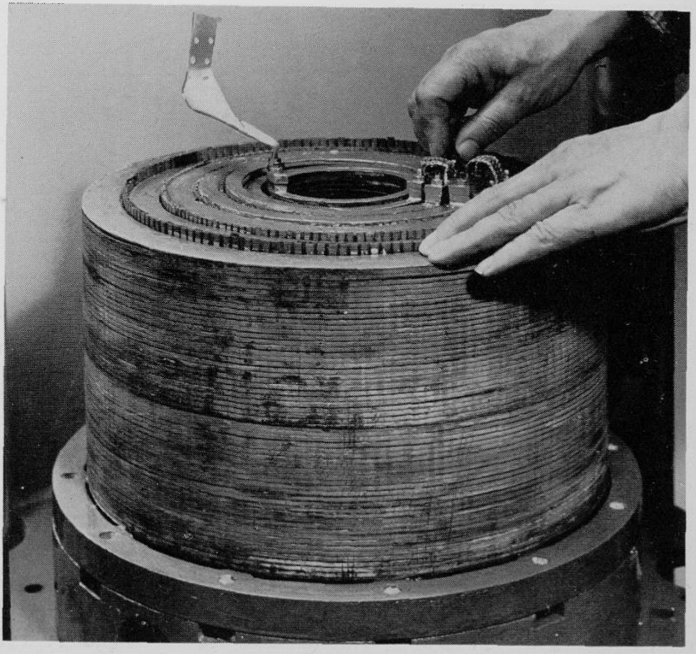


Figure 47
The strip-wound solenoid designed by M. Wood.

Disc constructions have been very successful. That due to F. Bitter working at MIT (figure 48) consists of a coil, about the size of a car wheel, built up from a series of slotted perforated copper discs interleaved with thin insulating discs. The perforations are uniformly distributed and in an array which is repeated at, say, 20° intervals round the disc. The insulating discs each have a 20° segment missing. By lining up the edge of a missing segment in an insulating disc with the slot in a copper disc a contact area is exposed on the latter which connects with the next copper disc. When a stack of discs and insulators is put together it forms a continuous conducting helix. This construction has been used at the Naval Research Laboratory in Washington to generate fields up to 15.8 T and at Malvern to give fields of about 13.5 T.

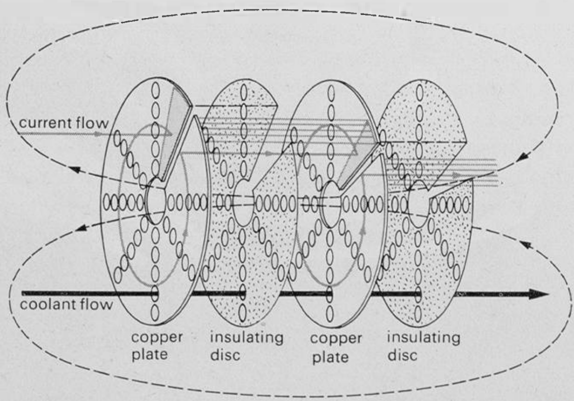


Figure 48
The disc solenoid developed by F. Bitter.

The 20 T magnet used by the National Magnet Laboratory at MIT is also of the disc type and was designed by D. Bruce Montgomery. The cooling channels are radial slots – like wheel spokes – chemically etched into copper discs. The discs are interleaved and insulated as before. The magnet has a 66 mm diameter bore and at full power consumes about 16 MW and requires 9 m^3 of cooling water per minute. Fields as high as 25.5 T have been generated by introducing iron poles along the axis of the coil and holding them a short distance apart. The use of such poles leaves a very small working space which is, nevertheless, sufficient for simple room temperature work.

Superconducting solenoids

The problem of how to reduce the power consumed in a solenoid and so also reduce the cooling required can be attacked in another and very effective way. Use can be made of the phenomenon of superconductivity – the lack of resistance to an electric current exhibited by certain substances when cooled to temperatures near absolute zero. In principle, fields can be generated in superconducting solenoids without any power dissipation.

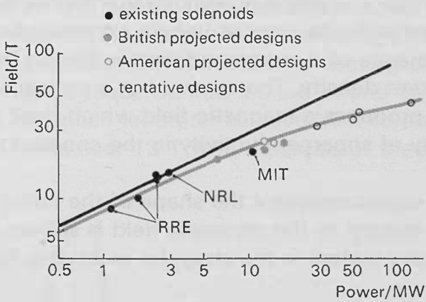


Figure 49
Characteristics of existing and speculative solenoids are shown in the graph. The upper line represents optimized relationship between field and power. Points on the lower part of the lower curve correspond to published performances of solenoids, while points on the upper part show the performance with current distribution chosen to overcome the strength problem.

Superconductivity was first observed in 1911 by K. Onnes but, until about 1960, physicists had thought that this property would be destroyed by the application of a relatively small external magnetic field and that it would be useless for generating high magnetic fields. The very presence of any sizeable current in the superconductor would give rise to a magnetic field which would, in turn, destroy the superconductivity. However, J. E. Kunzler and his colleagues at the Bell Telephone Laboratories found in 1961 that certain metallic materials remain superconducting even in intense fields. These alloys – known as hard superconductors – can be divided into two types: a set of ductile alloys of which the niobium–zirconium materials are probably the most useful and a set of very hard and relatively intractable compounds of which niobium–tin (Nb_3Sn) shows the most promise at the moment.

In a superconducting material the critical magnetic field – the field required to destroy superconductivity – varies with temperature. Extrapolating the early experimental results towards absolute zero, this critical field was estimated to approach 50 T. However, in practice, the upper limit seems to be about 25 T – and it is unlikely that higher fields will be produced with the present known superconductors. The performance of a superconductor is limited by the critical current density. The current has a critical value because it produces a magnetic field which itself destroys the property of superconductivity in the conductor.

For a hard superconductor the shape of the curve relating the critical current to the magnetic field is shown in figure 50. The important feature is the shoulder extending to high fields.

Whereas the alloy niobium–zirconium can carry current densities as high as 10^9 A m^{-2} in fields of 6–7 T, niobium–titanium will carry a lower current density in a field approaching 12.5 T. The titanium alloy can, therefore, be used for the inner windings of a solenoid of which the outer could be of the zirconium alloy. Such coils can produce fields above 10 T.

Nb_3Sn is the most promising compound because, at about 9 T, the critical current density is well above 10^9 A m^{-2}. Several methods of overcoming its intractability have been used so that wires or strip conductors can be made. In the earliest but still successful method a thick walled niobium tube is filled with a mixture of powdered niobium and tin in the right chemical proportions. The ends are sealed and it is drawn into a composite 'wire' and wound into a solenoid. It is then sintered, at about 1000 °C for 16 hours, to produce a wire consisting of an Nb_3Sn core with a niobium sheath.

An alternative method is to deposit thin layers of Nb_3Sn by decomposing suitable halide vapours (usually the chlorides) on to the surface of thin strips of niobium or even stainless steel as substrates. The substrates are usually about 0.03 mm thick and of almost any desired width up to several centimetres. The layers of niobium–tin are about 0.01 mm thick. Such a strip 5 mm wide can carry currents of several hundred amperes. The strips can be produced in lengths of more than 1000 metres and can be wound into solenoids after preparation. Yet another approach is to cover the outer surface of a niobium wire with tin by a dipping procedure and produce the compound on its surface by heating afterwards. Such wires may be made up into a multi-strand conductor. At least two American laboratories have built solenoids which can produce fields near 13 T with materials of these kinds.

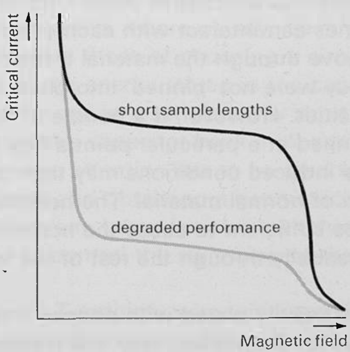


Figure 50
Critical current curve against magnetic field for a hard superconductor. The upper curve indicates the shoulder expected with short sample lengths of wire. Degradations in current density due to 'flux jumping' become important when long lengths of wire are used (lower curve).

Superconducting solenoids produce a number of practical problems. One fairly obvious one arises from the energy stored in the magnetic field. At 10 T this is about 35 MJ m^{-3} and rises as the square of the field. In the event of the coil being overloaded, it will become a normal conductor at a particular point at which the magnetic energy will be dissipated as heat while the super-current decays. The sudden release of considerable quantities of heat at liquid helium temperatures is undesirable. Even in small solenoids the quantities would be sufficient to fuse the wire and destroy the windings locally.

However, quite the most severe problem arises – particularly with the niobium–zirconium alloys – because long lengths of wire wound into solenoids have not shown characteristics as favourable as those found from short samples. Apparent degradations of the current density by factors as great as five have been reported. This is believed to be due to 'flux jumping'. In a strong magnetic field these superconducting materials are in a state in which the magnetic field has penetrated right into them. Owing to the natural inhomogeneity of the material the lines of magnetic flux are inevitably forced into bundles of some kind. Now the current and the flux lines can interact with each other. The flux lines would move through the material if their motion was not impeded or they were not 'pinned' into place by the presence of inhomogeneities. However, if a bundle of flux lines becomes unpinned at a particular point a flux jump occurs and the locally induced conditions may then be sufficient to create a region of normal material. The heat developed at this region could be sufficient to cause the normal region to be propagated thermally through the rest of the wire.

If the wires are heavily plated with copper, the copper sheath provides an alternative, very low resistance path for the current. The normal resistivity of a hard superconductor is many times that of pure copper at helium temperatures. In this way thermal propagation of a normal region can be prevented. Techniques of this kind can lead to a solenoid performance near to that predicted by short sample tests.

As has already been said, fields well in excess of 10 T have already been produced with superconducting solenoids. We can anticipate fields near 20 T being produced during the next few years.

Pulsed fields

By generating fields in extremely short pulses the problem of cooling can be ignored. Pulsed fields represent the only practical way of generating magnetic fields above 45 T and, as yet, fields above 25 T can only be reached in this way.

As early as 1924 P. Kapitza, then working at the Mond Laboratory, Cambridge, succeeded in producing pulsed magnetic fields in excess of 20 T. He did this by short-circuiting a generator for a period of milliseconds through a low resistance coil in which the field was to be produced.

However, during the past 20 years the simple capacitor discharge method of producing pulsed fields has come to be the most widely used. A large capacitor bank of some thousands of microfarads (μF) capacitance is charged to a voltage which may range from a few kilovolts to tens of kilovolts. The system is discharged through a small copper solenoid which may have a bore of, say, 10 mm. The practical problems are those of choosing a circuit so that a high proportion of the energy stored in the capacitors can be dissipated in the coil in a satisfactorily short time interval. The larger the capacitor bank in physical dimensions the more attention has to be paid to the inductance of leads and obtaining a geometrically balanced interconnecting system. Large ignitron switches (gas-filled valves) are used to control the switching of the circuit.

For fields below 25 T ordinary wire windings can be used but above this value disc constructions become necessary. These are similar to the Bitter solenoids but without the cooling holes. A 4000 μF system charged to 2.5 kilovolts and discharged through a coil cooled to liquid nitrogen temperatures (77 K) is a typical system. In a 6 mm bore disc type solenoid fields of about 41 T can be generated in overall pulse lengths of 0.25 milliseconds. In a 12.5 mm bore wire-wound system the field can rise to 25 T in an overall pulse length of 6 milliseconds. A typical repetition rate is about six cycles per minute.

To push the field to higher values correspondingly shorter pulse lengths must be used. Thus, to contain fields near 100 T the pulse length has to be shortened to about ten microseconds.

Explosive techniques have to be employed to extend the available range beyond this value. A strong magnetic field – about 40–50 T – is first generated in a copper cylinder by a

Figure 51
High explosives were used in this experiment at Los Alamos to create a field of more than one thousand tesla for a period of one microsecond. This photograph was taken a fraction of a second after detonation. Part of the firing bunker is visible.

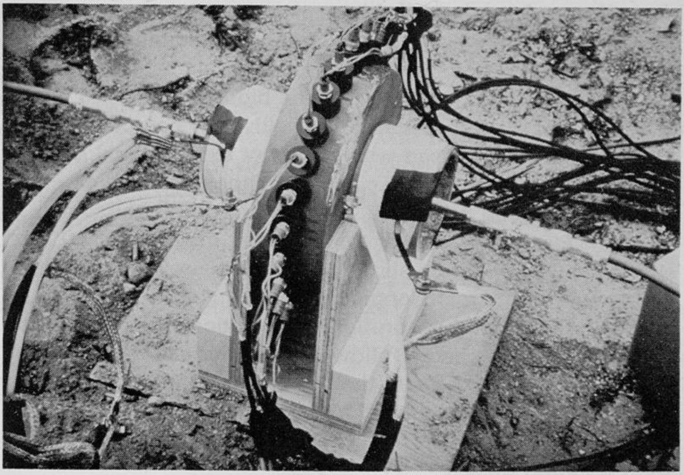


Figure 52
Implosion set-up at Los Alamos includes a cylindrical explosive charge with a ring of detonators glued to its outer surface and two magnetic pick-up probes. Not visible in the photograph are a pair of coils under the glass fibre tape and a thin-walled stainless steel cylinder under the charge.

coil energized by a discharge from a capacitor bank. The overall pulse length of the field is about 100 microseconds. When the magnetic flux in the cylinder reaches a maximum the cylinder is suddenly compressed by detonating an encircling charge of high explosive. The cylinder can be regarded as a conducting plastic or fluid cylinder. Continued implosion squeezes the cylinder tightly and, because the latter is conducting, the contained flux cannot escape through it. This process continues until the magnetic pressure inside the cylinder counterbalances the implosive forces. In this way, fields between 500–1000 T, lasting for about a microsecond, can be generated (figures 51 and 52).

Uses of intense fields

These different methods of generating intense magnetic fields are complementary to each other in many ways. Pulsed fields remain supreme for the highest fields and it is sometimes argued that at lower fields they can compete with steady fields if used with sufficient experimental skill. There are, however, whole classes of experiment which cannot be tackled using pulsed fields. Inevitably the time constant for the duration of the field is at most about 10^{-3} second and often much shorter; recording instruments must follow suit. In magneto-optical experiments in the infrared, where the change in absorption of an optical process due to the applied field is being followed, the physicist is usually limited by the weakness of the source or of the absorption process. There is thus a need for long integration times (of the order of minutes) which are quite incompatible with short duration pulses.

Superconducting solenoids are here to stay and are steadily taking over as the main means of generating fields between 1 and, at present, about 7 T. This range will continue upwards during the next few years. Nevertheless, the cost of solenoids for the 10–20 T range should not be under-estimated. It may well be more than £10 000 each for the upper reaches of this range. The most significant point resulting from the development of superconducting solenoids is that they now make practicable a whole series of devices depending on high magnetic fields for their operation.

The first of these was a tunable, fairly narrow band, infrared detector developed by E. H. Putley in 1963. In this device the sensitive element is a piece of indium antimonide which is cooled to liquid helium temperatures. As the magnetic field is increased from a fraction of a tesla the photoconductive response exhibits a pronounced peak which moves steadily from about 0.5 mm to shorter wavelengths. At a field of 1.4 T the peak is at 100 microns and at 8.5 T it moves to about 28 microns. The necessary magnetic fields can be provided ideally by using superconducting magnets which operate in the same liquid helium bath as the detector itself.

Another device requiring fields of 5 T upward was described by I. B. Bott last year. It is a tunable continuous millimetre wave generator, capable of producing very narrow band coherent radiation in the wavelength range from about 4 mm to 1 mm at power levels of at least one watt. The radiation wavelength is directly dependent on the magnetic field. Roughly 2.5 T gives a wavelength of 4 mm and 10 T gives 1 mm. Shorter wavelengths are possible and indeed have been generated. They depend solely on the production of high enough fields. Superconducting solenoids make such a device both practicable and portable.

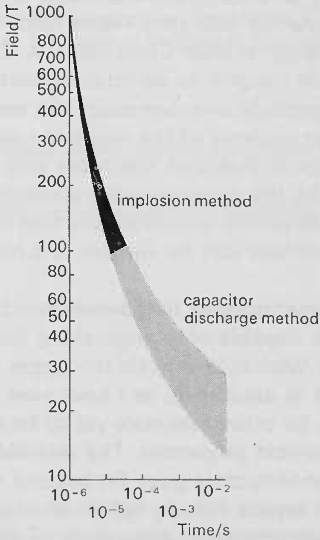


Figure 53
Performance of pulsed magnets in terms of field strength and time. The shaded area includes the work of groups based in the United States, Italy, and the Soviet Union. The volumes available for experiment vary from a few cubic millimetres at the shortest times, to tens of thousands of cubic millimetres for the longest.

The interest in producing high steady fields stems largely from solid state physics. Historically an early need was to generate very low temperatures close to absolute zero. To do this, powerful steady fields were required. For a time in the 1930s and late 1940s low temperature laboratories were the scene of much of the work on the techniques required for generating powerful magnetic fields. However, the upsurge in activities in solid state physics over the last ten years has produced many other interests. For example, with magnetically ordered materials, such as ferromagnetics, antiferromagnetics, and ferrimagnetics, the interaction of the magnetic moments of the individual ions in the material with the internal field (the Weiss field) is of the order of the thermal energy at their Curie point. A Curie point of 90 K corresponds roughly to an internal field of 45 T. With fields of this magnitude one can study the magnetic structure, or the ordered patterns of the magnetic moments of the ions, in the material. Fields of this order also approach that produced by the circumnuclear electrons on the nucleus of the simplest atoms, which means that the magnetic interaction with the nucleus can be studied in a new direct way.

Hard superconductors themselves need to be studied in fields quite capable of extinguishing their superconducting properties. With niobium–tin the upper limit to the critical field is still in doubt but, as I have said, is around 25 T. There may be other materials yet to be investigated with even more favourable properties. The technological importance of hard superconductors goes far beyond the device or instrument aspect already lightly touched on. The magnetohydrodynamic generation of electricity becomes an economic possibility if the necessary magnetic fields can be generated with superconductors.

The future

In conclusion, the urge to produce very high magnetic fields comes from the needs of research on the one hand and from the needs of technology on the other. In the next decade the upper limits for both steady and pulsed fields will be raised. With steady fields this will probably be by using a combination of superconductors to give a large region of intermediate field of say 15 T. An efficient water-cooled or even hydrogen-cooled system could then be used inside it to raise the field towards 40 T or perhaps beyond. The rewards for so doing are likely to be considerable, but whether it is feasible to progress beyond about 45 T remains a speculation. In the case of pulsed fields the technical difficulties beyond about 1000–1500 T are likely to be very great. How far beyond this we can penetrate is a question best referred to the future.

Dr David H. Parkinson is Head of the Physics and Electronics Department at the Ministry of Aviation Supply, Royal Radar Establishment, Great Malvern. He works in the field of solid state physics and its applications.

Acknowledgements

Figures 51 and 52: Los Alamos Scientific Laboratory, University of California.

Figure 47: Martin Wood, Clarendon Laboratory, University of Oxford.

Editor **Jon Ogborn**

Contents

The eight articles from *Science journal* collected in this book cover subjects ranging from the advanced techniques of building 'super cities' to the generation of ultrahigh magnetic fields, demonstrating some of the varied fields of technology in which physical principles are used. Since reprints from *Science journal*, now discontinued as a separate magazine, can no longer be obtained, this book has been produced to make a valuable source of supplementary reading material available to students of Nuffield Advanced Physics. It will also be of value to anyone who is interested in how engineers think.

United Kingdom 70p
New Zealand $2.35
Canada $2.95
Australia $2.35
(recommended)

Published by Penguin Education

ISBN 0 14
082.723 4