The Journal of

THE BRITISH INSTITUTION OF RADIO ENGINEERS

FOUNDED 1925 INCORPORATED 1932

"To promote the advancement of radio, electronics and kindred subjects by the exchange of information in these branches of engineering."

VOLUME 17 MAY NUMBER 5

CONVENTION VENUES

APART from the meetings held at Bournemouth in 1947 and at the 1951 Radio Show, all the Institution's post-war Conventions have taken place in Universities.

There is much to commend such a practice, not the least being the quiet atmosphere and available facilities which are more suited to the nature of Institution Conventions than those provided by hotels.

Nevertheless, there are many problems peculiar to the holding of such meetings within the precincts of a university, the main being the difficulty of accommodating all delegates in one college. This year, it is physically impossible to accommodate all delegates in King's College—the headquarters of the Convention. Indeed, almost as many delegates as are residing in the College will have to be accommodated elsewhere.

Although these circumstances emphasize the success of and need for the type of Convention which is being arranged, it is with regret that the Institution has had to announce that *no further reservations can be accepted.* The decision was prompted not only by lack of accommodation facilities; there is also the problem of the capacity of the lecture theatres.

In this latter connection, the Convention Committee has considered the alternative of arranging overflow meetings in adjoining theatres for all the six sessions. This would have many drawbacks, however, including the lack of personal contact with authors and discussion limitations, which would not be entirely overcome by the use of a television link.

It has therefore been decided that on each day two sessions will run simultaneously, i.e. sessions 1 and 2, 3 and 4, etc. The presentation of papers is being so organized as to give the widest possible opportunity for delegates to select the particular papers they wish to hear in all sessions; it will be easy for delegates to switch their attentions from paper to paper, or at any time during the three intervals which will occur during the day.

Should the Institution continue to hold its Conventions in universities? It may be argued that these various problems emphasize the necessity for the Institution to acquire as quickly as possible its own building. The urgency for such a building exists, if only for the purpose of enabling the Institution to hold more frequent meetings during the winter sessions, as well as other informal meetings. Adequate accommodation would also enable the Institution to hold at least an occasional Convention within its own precincts.

There is, however, much to be said in favour of continuing to hold Conventions in universities; whilst the meetings last only a few days, there is much to enjoy in an atmosphere of fundamental discovery and teaching, which enable delegates to realize a little more clearly that what is recently acquired knowledge will soon be history.

Overseas delegates particularly enjoy the privilege of such accommodation, and we might indeed say that in such ancient surroundings we have an excellent example of retaining what is good from the past and yet causing it to fulfil a useful function in the present.

Surmounting the difficulties which are involved emphasizes the ready co-operation which universities do give to the dissemination of new ideas and knowledge. Certainly it would be impossible to overcome the difficulties of this year's Convention were it not for the invaluable help given by Professor N. F. Mott and Mr. E. H. Dibden (Secretary) of the Cavendish Laboratories, and Commander C. E. Sclater, Bursar of King's College.

"ELECTRONICS IN AUTOMATION"

Third List of Papers Selected for Presentation at the 1957 Convention

Previous lists were published in the March and April issues.

Session 2—MACHINE TOOL CONTROL

Chairman: E. E. Webster, M.Brit.I.R.E.

(Thursday, June 27th—9.15 a.m. to 12.30 p.m. and 2.30 p.m. to 5.30 p.m.)

Numerically Controlled Machine Tools for the Production Engineer

by O. S. Puckle, M.B.E.
(E.M.I. Electronics Ltd.)

The paper discusses electronic methods of meeting the requirements of the production engineer for the numerical control of machine tools. The properties of different systems and components are examined.

The Design of a Basic System of Position Control for the Traversing Tables of Machine Tools

by K. J. Coppin, B.Sc. (Ekco Electronics Ltd.)

Machines are classified as "Progressive" and "Non-progressive" according to whether cutting and traversing movements occur simultaneously or alternate. The design requirements for both classes are analysed. Design details are given of a position control system for use with "non-progressive" machines. The use of a lead-screw as the measuring element are discussed and ways of correction and accuracy employed is shown to depend on the nature of the operation. Finally the use of punched cards for programming a series of operations and the extension of the principle to include "progressive" machines are discussed.

Some Aspects of the Application of Closed Loop Servo Systems to Machine Tool Control

by R. J. F. Howard
(Lancashire Dynamo Electronic Products Ltd.)

The paper discusses the combined design of machine tool and control equipment with particular reference to power requirements of the feeds on a two-axis profile follower. The effects of "backlash" and "stiction" on performance are considered.

Session 3—CHEMICAL AND OTHER PROCESSES

Chairman: Denis Taylor, Ph.D., M.Sc., M.Brit.I.R.E.

(Friday, June 28th—9.15 a.m. to 12.30 p.m. and 2.30 p.m. to 5.30 p.m.)

Industrial Applications of A.C. Polarography

by R. L. Faircloth, B.Sc., and D. J. Ferrett, M.A., D.Phil. (Atomic Research Establishment)

Polarography is a technique of wide use in industrial analysis both for metals and non-metals at concentrations down to a few parts in 100,000. Recent a.c. polarographs have very greatly extended the range and usefulness of these techniques. Concentrations of 1 in 10,000,000 can be measured without the disadvantages of conventional polarographs. The principles of the square-wave polarograph are discussed and recent work on the pulse polarograph and the r.f. polarograph is described. The range and versatility of these instruments offer many opportunities for plant-control and instrumentation.

Approximate Relations between Transient and Frequency Response

by H. H. Rosenbrock, B.Sc.(Eng.), Ph.D.
(Costain-John Brown Ltd.)

An existing method for inter-converting the frequency and transient responses of a stable, linear device becomes inconvenient (*a*) when the frequency response of the device has a wide pass-band over which the phase changes greatly, (*b*) when the input is not a step-function but a suddenly-applied sinusoid, (*c*) when the transient response is known, and the frequency response is required for the whole range of frequencies over which its magnitude is appreciable. The paper shows how the existing method can be modified to avoid these difficulties. The results are used to evaluate the performance of a distillation column.

Session 4—AUTOMATIC MEASUREMENT AND INSPECTION

Chairman: J. E. Rhys-Jones, M.B.E., M.Brit.I.R.E.

(Friday, June 28th—9.15 a.m. to 12.30 p.m. and 2.30 p.m. to 5.30 p.m.)

Automatic Inspection as the Key Control Element in Full Automation

by John A. Sargrove and Denis L. Johnston, B.Sc.(Eng.) (Automation Consultants and Associates, Ltd.)

A fully automated process is essentially an exercise in "system design." The feedback signals to control and stabilize the system will be originated by inspection or measuring devices, dependent on the human senses or automatic instrumentation. The characteristics of a number of classes of such systems are distinguished, as are the special requirements of the inspection equipment, which must have a higher order of reliability than is usually associated with laboratory measuring instruments. A good visual display of the information derived, plus some degree of manual control, may often be a more economic solution than an attempt to achieve 100 per cent. automation.

Automatic pH Control

by R. S. Evans, M.A. (W. G. Pye Limited)

The paper will cover the field of pH and allied instrumentation, and its application to automatic process control. It explains the modern methods of measuring pH (from an electrical point of view) including a description of the measuring electrodes and requirements of the electronic circuits employed. It also describes four applications in the manufacture of paper, control of electroplating, treatment of water and effluent.

Session 5—SIMULATORS

Chairman: Professor D. G. Tucker, D.Sc., Ph.D., M.Brit.I.R.E.

(Saturday, June 29th—9.15 a.m. to 12.30 p.m. and 2.30 p.m. to 5.30 p.m.)

Computing Applications where Analogue Methods are Superior to Digital

by R. J. Gomperts, M.A. (English Electric Co. Ltd.)

In solving scientific and engineering problems and design studies the first step consists of formulating the mathematical analogue of the problem which is to be studied. It is then a a matter of convenience which means are employed to realize this mathematical model—analogue or digital. There are therefore no *a priori* reasons why either method is to be preferred and the choice will be dictated by the exigencies of the problem and the equipment in existence. It is, therefore, not surprising that for, many problems, analogue machines are quoted and three particular problems are studied for which analogue methods appear to be superior.

Electronic Synthesis of Flexible Beam Behaviour

by M. Squires, B.Sc., and W. G. Hughes, B.Sc. (Royal Aircraft Establishment)

The paper gives an account of analogue computing techniques for handling the partial differential equations concerned in the dynamic behaviour of elongated flexible beams. These techniques are based substantially on the use of systems of difference equations to replace the original equations.

An electronic model depicting translational as well as rotational motion is first discussed. In a second model only rotational motion is allowed to appear explicitly in order to reduce the special difficulties associated with "free–free" beams.

An hypothesis concerning the mechanism of structural dissipation is advanced. Its advantages from the standpoint of computer dispositions are discussed in relation to its theoretical viability.

Analogue Computers and their use in Nuclear Reactor Safety Studies

by I. Wilson, B.Sc.(Eng.) and R. Potter, B.Sc. (Atomic Energy Research Establishment)

Analogue computers have become the principal means of obtaining solutions to the differential equations representing the transient behaviour of a reactor system. With suitable analogue computers it is possible to study complete nuclear plants including the reactor, heat exchanger and

Session 5 (*contd.*)

turbine. Safety studies include an examination of the overall stability of the system, the effects of coolant pump failure, burst steam lines and control rod maloperation. The paper describes computational and circuit techniques which have been used successfully at A.E.R.E., Harwell, to study the various aspects of nuclear plant kinetics which are relevant to reactor safety.

A Nuclear Power Plant Training Simulator for Use at Calder Hall

by I. Wilson, B.Sc.(Eng.) and L. A. J. Lawrence (Atomic Energy Research Establishment)

In designing a simulator for training personnel in the operation of Nuclear Power Stations it is necessary to achieve a suitable compromise between realism on the one hand and size and cost on the other. This paper outlines the problems peculiar to nuclear reactor and heat exchanger simulation and techniques are described which deal satisfactorily with these problems. A description is given of a machine which gives a reasonably accurate simulation of the behaviour of stations of the Calder Hall type without being too complex in design.

A High Speed Analogue-to-Digital Converter

by G. J. Herring, M.Sc., and D. Lamb, M.Sc. (Royal Aircraft Establishment)

Some control and computing problems are discussed in which advantage may be gained by encoding analogue quantities into digital form for subsequent processing. A voltage-to-digital converter has been designed for use with an electronic analogue computer in order to combine analogue and digital computation techniques. The system is basically a servo with a digital-to-voltage converter as a non-linear feedback element. A binary register, driven from a 100 kc/s clock pulse, is operated by gates arranged to allow an increasing or decreasing count, dependent on the sign of the error between the voltage input to the converter and the voltage analogue of the number in the register. Functional descriptions of the main parts of the equipment are given, with reference to some of the factors that have been found to be critical. The apparatus is now working, but experience has shown that considerable increase in both speed and accuracy is possible, and modifications which would provide this improvement are described.

Session 6—AUTOMATION IN THE ELECTRONICS INDUSTRY

Chairman: L. H. Bedford, C.B.E., M.A., B.Sc., F.C.G.I., M.Brit.I.R.E.

(Saturday, June 29th—9.15 a.m. to 12.30 p.m. and 2.30 p.m. to 5.30 p.m.)

Recent Advances in Automatic Component Assemblies

by K. M. McKee, B.Sc. (Geo. Tucker Eyelet Co. Ltd.) and L. H. Gipps (British United Shoe Machinery Co., Ltd.)

Reference is made to recent and probable future development of machines for the preparation of electronic components and their automatic assembly into printed wiring boards. The design of components and the processing of printed wiring boards is discussed in relation to the requirements of automatic assembly and mention is made of current practice with regard to the problem of interconnection and the incorporation of sub-assemblies. The field of application of fully automatic machines and that of semi-automatic machines using manually operated bench machines is surveyed in the light of production experienced gained up to the present date concluding with some indication of possible future trends in mechanized production.

Automatic Methods in Radio Manufacturing Processes

by D. Stevenson and R. B. Shepherd, B.Sc.(Eng.) (Kolster-Brandes Limited)

The first part of the paper is concerned with the design and development of a novel electro-magnetic sensing device with its associated amplifier and switching circuits for interlocking power presses converted to automatic operation.

The second part describes the design and development of automatic coil winding machines. Details are given of the electronic binary counter and the methods of preset switching.

The final selection of papers for the six sessions of the Convention will be given in the Convention Handbook. Copies will be available early in June, and will be sent to all persons registering to attend the Convention.

THE DEVELOPMENT OF NUCLEAR POWER*

by

Sir John Cockcroft, O.M., F.R.S.†

The 1957 Boyle Lecture of the Oxford University Scientific Club. Delivered on 5th March 1957.

SUMMARY

The Calder Hall type of nuclear power station is described and its future development discussed. Other types of reactors are mentioned in connection with various applications. The power requirements of Commonwealth and European countries reviewed. Current researches on the fusion reaction are described briefly.

1. The Calder Hall Power Station

The development of nuclear energy for the production of large-scale electric power reached an important stage when the Calder Hall nuclear power station came into commission in May of last year and was formally inaugurated by H.M. the Queen in October.

This power station derives its heat from two nuclear reactors which are the furnaces of the nuclear power station. The heat is then used to raise steam and drive conventional turbo-generators. So the essentially new part of the power station is the reactor. A diagrammatic view is given in Fig. 1.

The reactor consists of a core of 1,200 tons of graphite blocks contained in a pressure drum about 40 ft in diameter. In channels between the graphite blocks are stacked, one above the other, 10,000 uranium fuel elements 4 ft long by 1·15 in. in diameter. These are sheathed in Magnox—an aluminium alloy, and the sheath has transverse fins to improve the heat dissipation.

The nuclear chain reaction develops when the critical quantity of uranium fuel elements is loaded. The chain reaction is thereafter controlled by control rods which move into the reactor core and damp down the chain reaction as they move in. The full charge of uranium metal is about 130 tons, and the heat equivalent of 180 electrical MW is generated at full power. The heat is transferred from the hot uranium rods to four steam generators by circulating carbon dioxide gas under a pressure of 7 atmospheres. The gas leaves the reactor at a temperature of 336° C and flows past tubes full of water in the steam generators. Steam is then produced and goes into high pressure and low pressure drums at temperatures of 590° F and 340° F respectively. These steam conditions are poor by comparison with the most modern power stations. This results from the temperature of the fuel elements being limited to about 400° C in the first station. The steam flows to four 23 MW turbo-generators. About 15 MW of these are used internally to drive circulating fans for the carbon dioxide and for other auxiliaries, and about 70 MW flow out to the electricity grid.

The uranium fuel rods are charged into the reactor and discharged through channels in the top of the reactor and in the concrete shield above. The fuel elements would normally stay in the reactor for periods of over a year. They may, however, develop faults in the covering magnox sheath due to faulty welds or pin holes or distortion of the uranium leading to failure, by cracking. In this case the hot carbon dioxide gas would slowly attack the uranium, forming oxide, and radioactive fission products would begin to percolate into the coolant. This is at once detected by tubes which "sniff" the gas from each channel and have in their circuit detectors which respond to radioactive fission products, particularly the gaseous ones. So the leak is recorded before it is serious and the channel containing the

* Published by permission of the Oxford University Scientific Club. (Paper No. 395.)

† Director, Atomic Energy Research Establishment, Harwell.

U.D.C. No. 621.039.

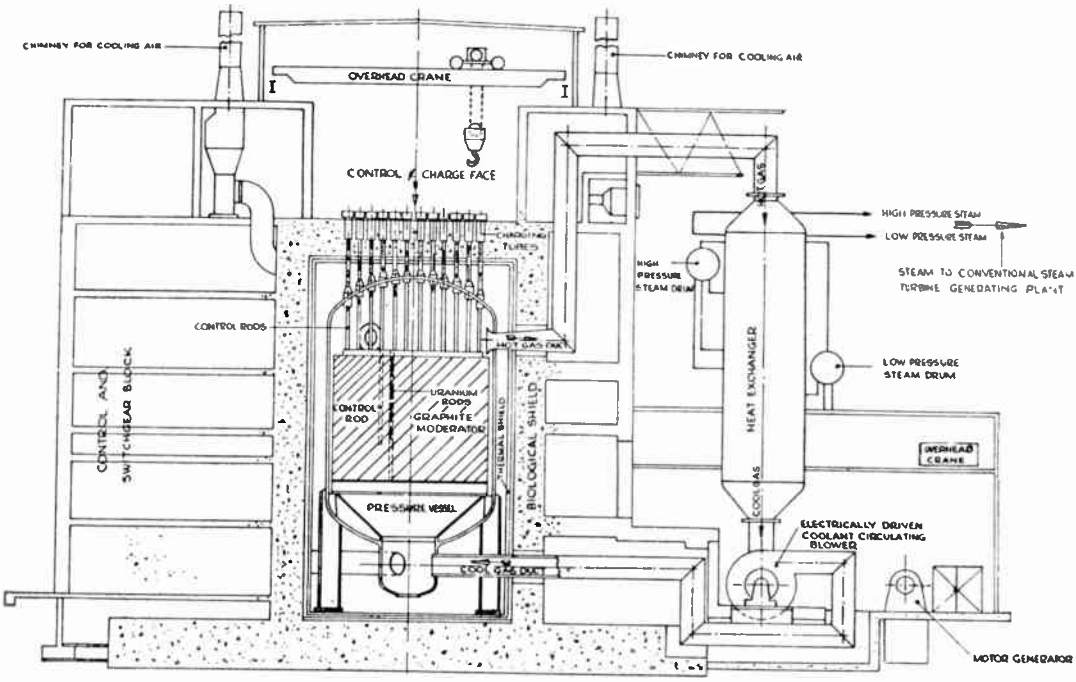


Fig. 1. Diagrammatic representation of the Calder Hall gas-cooled reactor.

defective fuel element can be discharged. During the first six months of operation, only two faulty elements in 10,000 were found and the faults could hardly be seen on inspection.

The main fault occurring in the fuel elements was due to a new effect, predicted by Dr. Cottrell, which is now known as the "Cottrell effect." Dr. Cottrell predicted that when uranium is traversed by the fission fragments—which are high speed heavy atoms, it leads to the production of local stresses in the uranium crystallites. As a result of this, if the uranium bar is under load, at a temperature of 100° C, creep will occur at stresses 100 times less than are required to produce the same creep rate normally. This was verified by making springs of uranium metal and loading them in the BEPO reactor. The springs were found to extend much more under irradiation. Since the bottom fuel elements of the reactor are stressed by the load of all those above, they tend to bow like a strut. The normal thermal creep would not have been troublesome; the creep induced by radiation, however, is larger, and we have verified on shutting down the first reactor that the Cottrell theory is largely correct. This can be remedied in the Calder Hall reactors by fixing intermediate supports to the fuel elements, so that the effective strut length **is** reduced. In future designs, the weight of the fuel elements will be carried by the graphite.

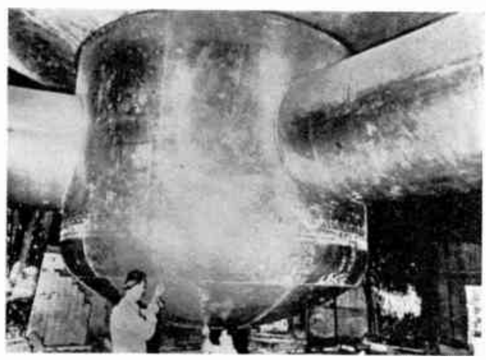


Fig. 2. View of the lower end of the pressure vessel during construction.

Another teething trouble has been a rather excessive leakage of the carbon dioxide gas which is used under pressure for cooling. At present about two tons a day is lost from valve spindles, glands and joints. This leakage is being progressively reduced by conventional engineering methods. It is not serious from a cost point of view. It has emphasized the importance of the most rigorous standards in testing the plant for leak tightness—the standards used in high vacuum equipment are required and can be achieved. Apart from these and other minor troubles the power station has been put into commission with far less trouble than a conventional plant, and has so far operated for 90 per cent. of the possible working hours. Normally power stations only operate for up to 80 per cent. of the full working hours owing to variability of the load, and we had expected this power station to operate for about 85 per cent. of the time.

Three other power stations similar to Calder Hall are at present under construction and the four will ultimately generate about 300 MW for the grid.

These first stations are to be followed by three power stations of much larger output. They are being built by industry. As a result of the experience of building Calder Hall it has been found possible to increase the size of the reactor containing vessel. Two of the designs will use spheres of about 70 ft in diameter. The third will use a cylinder of 50 ft in diameter, compared with the 37 ft of Calder Hall. As a result of this, the amount of uranium metal in the core has been almost doubled and will be 500 tons in one of the designs. The extraction of heat from each ton of uranium has been about doubled by increasing the pressure of the heat transfer gas about 50 per cent., and by improving the shape

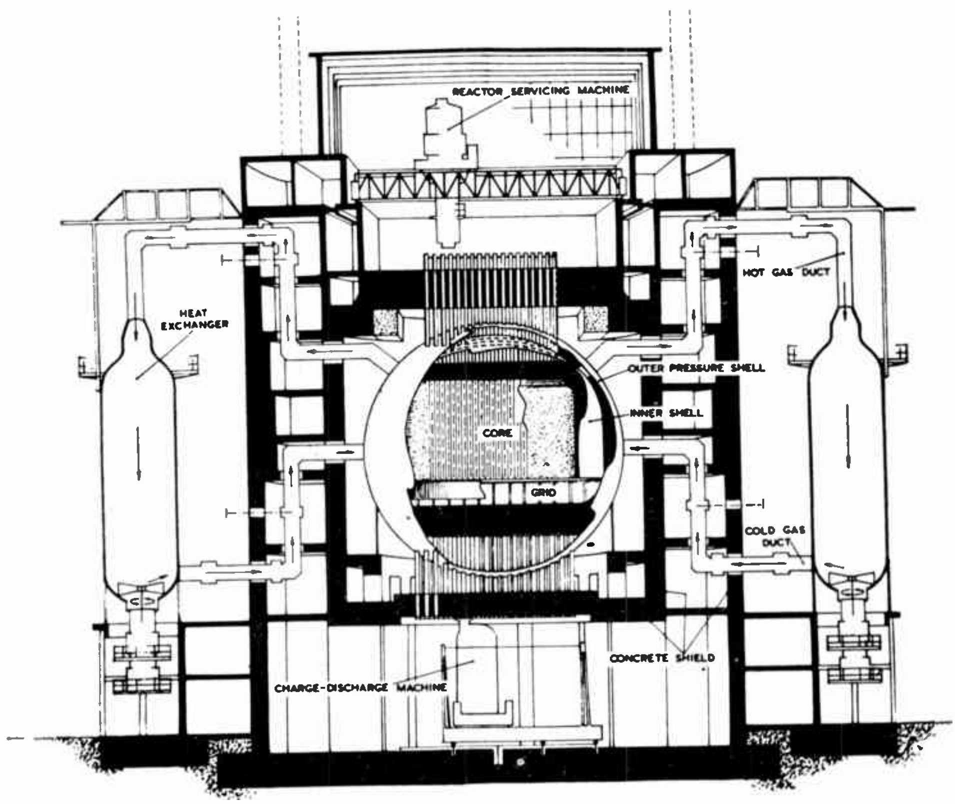


Fig. 3. The G.E.C.-Simon Carves reactor system. Most noteworthy departure from existing practice is the bottom charge and discharge facility.

of the fuel element surface. So by a combination of these causes the gross electrical output of the most powerful station has been increased three to four times over Calder Hall.

The cost of the stations has not yet been officially announced, but it has been stated in the technical press that the cost of the Scottish station will be between £35M and £40M for a net output of 300 MW guaranteed. So the capital costs will be between two and two and a half times the capital cost of a large-scale coal-fired station of the 1960's. Half the capital cost is due to the reactor and the remainder to the heat exchangers, turbo-alternators and site costs.

From each ton of uranium we expect to extract the heat equivalent of 10,000 tons of coal worth £50,000 at present prices. Uranium fuel costs are expected to be less than half those of coal-fired stations. The overall result is that nuclear power costs are likely to be slightly higher than costs of power from coal stations at 75 per cent. load factor, provided the price of coal or oil does not continue to rise faster than other commodities. This last assumption sems very dubious.

A more important point for this country is that our annual energy requirements are likely to increase by about 80 million tons a year coal equivalent by 1965. Annual coal production has been steady for four years but is planned to increase by 20 million tons by 1965. The remainder of the increase will have to come from additional oil or coal imports or from nuclear energy. Nuclear energy is therefore welcome as a method of reducing our reliance on oil from politically disturbed areas, and as a means of reducing the great future drain on our balance of payments which excessive fuel imports would lead to. We already import a net quantity of mineral oils and lubricants worth £270M a year.

If we look further ahead we can foresee a continued improvement in the performance and output of our gas cooled graphite moderated reactors. In particular, fuel temperatures are likely to be increased and so thermal efficiency and output will increase. So it is not at all unlikely that we shall see power stations with electrical outputs of 800 and 1,000 MW. These very large outputs will have the advantage of reducing the number of sites which will be required. Because of the large amount of cooling water which will be required, the sites will be limited mainly to coastal and estuary sites. But this would be so even for large coal-fired stations of the future.

2. Other Types of Power Reactor

Other types of nuclear power stations are

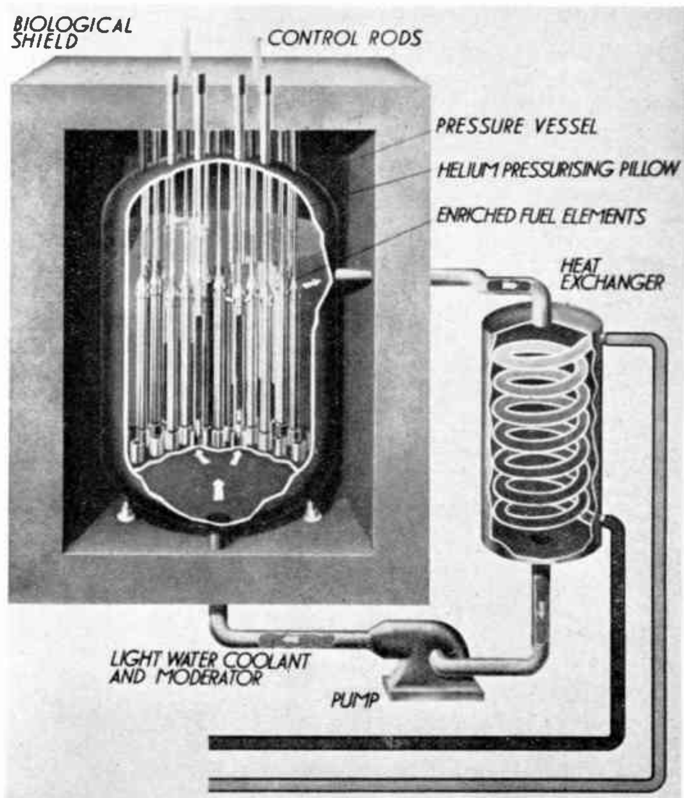


Fig. 4. A light water-moderated enriched fuel reactor.

possible. The U.S. is putting most of its effort and money into the development of water moderated reactors, in which ordinary water surrounds fuel elements consisting of uranium oxide sheathed in an alloy of zirconium. Such reactors will only react if the uranium fuel is "enriched." The proportion of U235 must be about 50 per cent. above the normal content. So the uranium fuel costs rather more than three times the cost of natural uranium. The cost of sheathing in zirconium alloy is also high. So fuel costs are at present very high—about equal to our total power costs at the present time. They are, however, expected to be reduced by metallurgical development. The capital costs of this kind of nuclear power station are likely to be appreciably lower than those of our first electricity authority power stations, because they are smaller; the reactor drum is 10 ft in diameter instead of the 35 ft of Calder Hall.

At Harwell we are working on more advanced types of nuclear power stations. One of these aims at working at appreciably higher fuel temperatures—of the order of 800° C—by using ceramic fuels. It would also provide inherently higher outputs per unit volume, which is good for reducing capital costs. We are at present working on its technology, which is an indispensable preliminary to all reactor development. For example, samples of the fuel elements have to be tested in so-called "loops" in our research reactors BEPO and DIDO. These loops are complicated test equipment. They contain the fuel element samples, and past them circulates gas of the same kind and at the same temperature as that to be used in the final reactor. The loops are run continuously for months, and then the fuel element is removed in a highly radioactive state and examined behind lead walls by remote handling equipment. Simultaneously with operating loops, we design the first reactor experiment, and this normally takes about 2½ years to design and build. The reactor experiment allows us to obtain experience of all the operating problems of new types of reactor and the nuclear physics behaviour and the kinetic behaviour of the reactor. If the reactor experiment is promising, the next step is to design and build a prototype reactor, which is probably a small scale version of the later full scale commercial models.

We are also building an experimental fast neutron breeder reactor with the long term objective of making much greater use of our uranium fuel supplies. Our first electricity authority reactors will extract the heat equivalent of 10,000 tons of coal for each ton of uranium. In principle however it is possible for breeder reactors to do 100 times better, because the breeder can burn the abundant U238 or thorium. We started this development by building two so-called "zero energy" reactors, ZEPHYR and ZEUS, fuelled by plutonium and U235 respectively. In the latter there is a core consisting of rods of U235 surrounded by a "blanket" of natural uranium. The chain reaction starts in the core and surplus neutrons are captured in the blanket to produce Pu239. Some of the fast neutrons also produce fission in the blanket. In the ZEPHYR reactor, which has plutonium fuel, more than two new fissile atoms are produced for each fissile atom destroyed. We say the "breeding gain factor" is over 2. The fact that the neutrons are slowed down as little as possible means that the capture and loss of neutrons by structural materials is small. In a full scale fast reactor fuelled by plutonium we might expect a breeding gain of 1·6 to 1·7.

The fast reactor is therefore the nuclear physicist's dream. On the other hand it presents severe problems to the engineer. For the core of the reactor is small, being fairly concentrated plutonium or U235 contained in a cylinder of about 2 ft diameter. From this small core the equivalent of about 60,000 kW of heat has to be extracted in our first experimental model which is now being built at Dounreay. The heat developed in the core is removed by circulating liquid sodium—chosen because it slows down neutrons very little. The liquid sodium passes through twelve heat exchangers arranged in parallel. Surrounding the whole reactor is a great sphere to provide a safeguard against any release of radioactivity. This reactor is due for completion in 1958. A schematic illustration is given in Fig. 5.

3. Nuclear Propulsion for Ships

The development of atomic energy for the propulsion of ships has already been shown to be technically feasible by the performance of the U.S. submarine *Nautilus*, which has proved to be reliable in service and has already steamed more than 50,000 miles. The costs of commercial propulsion using a Nautilus-type propulsion unit have however been stated to be about six times present costs for commercial

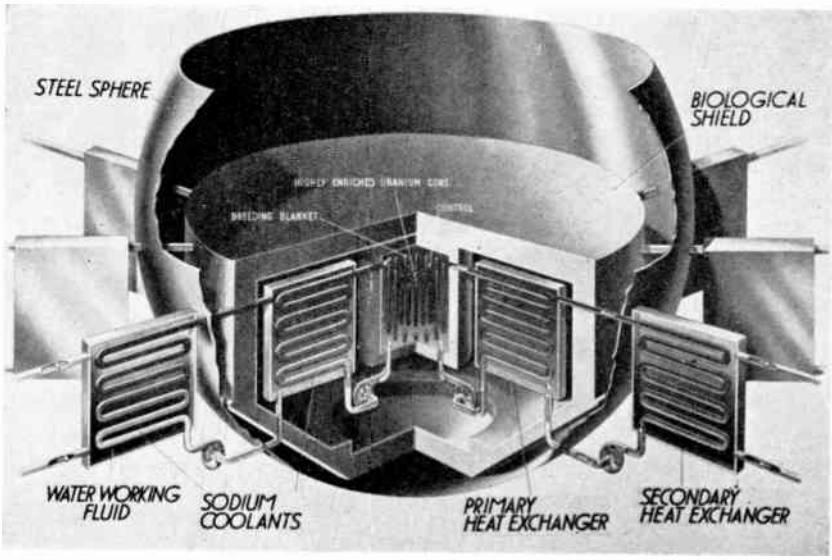


Fig. 5. Fast reactor employing liquid sodium heat transfer.

shipping. The United States is now building a combined passenger and cargo ship with a deadweight cargo capacity of 12,000 tons driven by a similar propulsion unit. The cost of construction has been estimated at $42M and this ship is not intended to be economic.

In Britain the possibilities of commercial propulsion are being investigated by a combined team of the British Shipbuilding Research Association and the Atomic Energy Authority. In order to make best use of the economic characteristics of nuclar power units, applications requiring high shaft horsepower and having a high effective utilization are obviously the first candidates. So a large high-speed tanker designed for the Cape run would be an interesting project to study—we have not yet embarked on this.

In this country we are studying the organic liquid moderated reactor for ship propulsion. If technically feasible it would have the advantage of lower fuel costs than the pressurized water reactor of the Nautilus type, since fuel corrosion problems would be less severe. There are other technical advantages such as a low system pressure and less radiation in the external circuit.

4. Overseas Developments

The application of nuclear power in other countries depends a great deal on their resources of conventional fuels. If we look at some of the other countries of the Commonwealth, we find that India is a "power hungry" country. At present India uses only about 1/10th of the energy per head that we use in Britain, and this is the same ratio as our income per head. India's electrical power consumption per head is only 1/80th that of Britain, but it is planned to increase this sevenfold by 1975, increasing from 3,500 MW to 25,000 MW. However, the potential hydro-electric power is 35,000 to 40,000 MW, so this would be far from exhausted by 1975 and hydro-electric power is generally much cheaper than power from coal and likely to remain cheaper for the next two decades. So in India, nuclear power will be introduced only in areas remote from hydro-electric stations. However a good beginning has been made in India by the founding of a nuclear research centre under the direction of an eminent theoretical physicist, Dr. Bhabha.

They have already built on an attractive site outside Bombay a small research reactor with some help from Harwell and are now building a more powerful reactor which has been presented by Canada, being a duplicate of the powerful heavy water reactor we built at Chalk River. So the Commonwealth is pulling well together.

In Australia there are ample sources of coal in New South Wales and there is a great hydro development in the Snowy Mountains. But South Australia has very limited coal resources and in Queensland there are important mining areas remote from the other energy resources. So nuclear power might well fill a place.

In Western Europe, a considerable amount of power is now generated from expensive American coal costing £8 to £10 per ton, or from imported oil. Over half of Sweden's energy is already supplied by imported fuel and this proportion is rapidly growing. The time is therefore approaching when countries like Holland, Belgium, France, Germany and Sweden will start building nuclear power stations and they might well make use of U.K. designs. If the Euratom plan develops, this might give a considerable acceleration to the programme.

In general therefore mid-1965 should see the world programme well launched; after that it should grow rapidly in importance and output.

5. The "Fusion Reaction"

I would like finally to look still further ahead and speak briefly about the possibility of using the reaction between the light elements—the so-called "fusion reaction" to produce power. The objective is to make use of the nuclear reactions between two deuterons or between a deuteron and a triton to release energy. You may remember that these nuclear interactions were discovered in the Cavendish laboratory during the golden age of nuclear physics of 1932-3.

It is well known that the stars derive most of their energy from the reactions between the light elements. As temperatures increase due to gravitational contraction, the velocities increase, and the penetration of the nuclei, can occur at increasing rates. The result is that the energy release increased. Also temperature must be very high for fusion reactions to go and particle densities must be reasonably high.

A fusion reactor must provide a source of energy to heat up the gas to the reacting temperature. Second, it must provide a means of containing the hot gas, so that losses of heat by radiation, conduction and convection do not far exceed the heat input. If we look at the radiation processes first and were to assume radiation equilibrium, then the usual black body relationship would give a radiation flux of 10^{21} watts/cm^2 at 10^8 degrees Kelvin. However, radiative equilibrium is far from achieved. A complete calculation of energy losses shows that since reaction rate increases very rapidly with temperature, and the radiation rate increases as $T^{\frac{1}{2}}$, power produced may exceed radiation above a certain temperature.

The important question is how we can both heat up and contain the hot reacting gas. Experimental work on this subject started in the Clarendon Laboratory in Oxford under Dr. Thonemann in 1948 and a little later in Imperial College under Sir George Thomson.

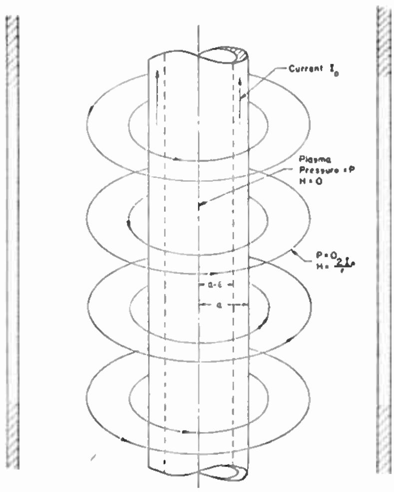


Fig. 6. Schematic representation of pinch effect in plasma.

The principle of one method is to heat up the gas by a powerful transient electrical discharge from capacitors, and to constrain the plasma by magnetic fields. Fortunately the experimenter is helped by the well-known pinch effect whereby the magnetic field of the current itself squeezes the discharge into a narrow filament. This can be helped by superimposed magnetic fields.

Experiments of this kind carried out in straight tubes were described by the Russian physicist Kurchatov at Harwell last May. A powerful discharge carrying a million amperes was produced by a capacitor bank. The discharge was observed to contract and expand again. The experimenter observed X rays of about half a million volt energy and neutrons were observed on the second contraction of the discharge channel. At first the Russians thought they were observing a true thermo-nuclear reaction but on further investigation they decided against this interpretation of the neutron production—for several reasons.

At the Clarendon Laboratory work also started on discharges in straight tubes, but later work, and all work at Harwell, has been concentrated on producing these pinched discharges in toroids since this avoids the

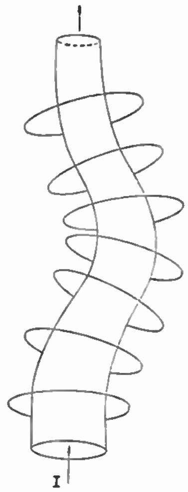


Fig. 7. "Kink" instability of the pinch effect.

troublesome problem of having electrodes which naturally get very hot under these powerful discharges. An account of these early researches has been published in a series of papers in the *Proceedings of the Physical Society* (B, January 1957).

One of the papers, by Carruthers and Davenport, describes one of the troublesome characteristics of these ring discharges. The discharges turn out to have a very powerful instability characterized by wriggling of the discharge in the confining torus. A very simple explanation of this is that if the filament turns itself into a helix the magnetic energy is increased. So it naturally strives to become a helix. Since the magnetic field lines crowd together on the concave side of the curved current channel, there is a greater magnetic pressure there and so the concavity increases. Fortunately there is an opposite effect, predicted also by classical physics, which helps. As the wriggling filament approaches the walls of the tube, eddy currents in the metal tube walls tend to repel the filament. These effects were investigated at Harwell in a glass torus of 10 cm bore and 76 cm mean diameter, which heated the core by a large pulse transformer. Currents of 1,500 amperes were produced in argon and xenon at pressures of 1 to 5×10^{-3} mm Hg. The discharges were observed with an electronically gated image convertor which enables photographs of 8 microseconds duration to be taken at any desired moment during the pulsation.

The experimental programme consisted of observing the phenomena in such ring discharges: in particular, measuring temperature as a function of current and pressure. Temperatures are measured by classical spectroscopy methods using the Doppler broadening to measure ion temperatures. Electron temperatures are measured by micro-wave techniques. This field of research is a fascinating example of the rapid and recent growth of magneto-hydrodynamics—the inter-actions between magnetic fields in the motion of matter—which is of great importance in stellar phenomena.

We are now at the same stage as fission research was in 1940 when the possibility of a chain reaction was well understood but many uncertainties remained, and a further 15 years elapsed before large scale nuclear power was developed. It is certain, however, that there is a fascinating field of research ahead.

NOTICES

OBITUARY

The Council regrets to record the death on May 10th of Ernest John Emery (Member).

Born in October, 1899, Mr. Emery joined the Marconi Company in 1916 as a Wireless Operator. When radio broadcasting started in 1922, Mr. Emery transferred to the Marconiphone Company; he was Service Manager at the time of the incorporation of the Company into the Electric and Musical Industries Group, and subsequently held many responsible positions within the Group. He was appointed a Director of the main Board in February, 1956.

Mr. Emery's association with the Institution started in 1930. He was elected a full Member in September, 1938, and had been a Trustee of the Institution's Benevolent Fund since 1952.

The Institution's work in initiating and promoting training schemes for technicians and service mechanics especially interested Mr. Emery. He assisted in the formation of the Radio Trades Examination Board, and officially represented the Radio Industry Council on the Board. Mr. Emery also served on a number of the Board's Committees and played a very important part in securing incorporation. He served as Chairman of the Board from 1951 until his death. Mr. Emery was also Chairman of the City and Guilds Advisory Committee on Radio and Television Service Work.

Insistent on the most meticulous presentation of detail, Mr. Emery invariably succeeded in carrying his colleagues with him, whether it be on a Committee or on any matter of negotiation. This ability often disguised a very generous disposition, a characteristic known only to friends with whom he associated in charitable matters.

New City and Guild s Institute Building

A new headquarters is to be built by the City and Guilds of London Institute at 76-78 Portland Place, London, to house its various administrative departments, which are at present scattered in a number of buildings in different parts of London. The new building will also considerably ease the problem of holding the many committee meetings associated with the Institute's work. It is expected to be completed in two years' time.

Members may be interested to know that at one time the Institution's Building Committee discussed with the ground landlords the acquisition of this particular site as a possible place for the new Brit.I.R.E. headquarters.

Graduate Course in Information Engineering

The Electrical Engineering Department of the University of Birmingham has established a twelve-month graduate course in Information Engineering. Intended for honours graduates in electrical engineering, the course leads to a M.Sc. degree.

Subjects studied will be five from the following: communication systems; communication and information theory; electronic computers; automatic control systems; advanced network theory; radio propagation and echo-ranging systems; electrical properties of materials. The first nine months will be spent in lectures, tutorials, and laboratory work, and the remaining three months on work on an individual project.

The fee for the course is £81 and further details may be obtained from the Supervisor of the Graduate Course (Dr. D. A. Bell).

Safety and Factory Efficiency Exhibition

The Safety and Factory Efficiency Exhibition, which is to be held in Birmingham from 14th-21st June, will emphasize the increasing applications of atomic energy in industry. The United Kingdom Atomic Energy Authority is collaborating in the provision of exhibits and also in the associated Congress, which is to be held from 14th-16th June. Speakers from the Authority will discuss: "Training for Atomics," "Isotopes in Industry," "Health and Safety when Using Radioactive Materials" and "Atomic Reactors."

The Exhibition and Congress, which is to be held at Bingley Hall, Birmingham, is organized by the Birmingham and District Industrial Safety Group; further information and applications for registration forms for the Congress may be obtained from Mr. W. G. Appleyard, Industrial Safety Training Centre, 22 Summer Rd., Acocks Green, Birmingham 27.

RADIO AND ELECTRONIC COMPONENT SHOW, 1957

THIS year the Radio and Electronic Component Manufacturers' Federation met the problem of the increasing number of exhibitors at the "Component Show" by taking extra space in another building, Park Lane House, a few hundred yards from Grosvenor House, London. A total of 164 firms took part and it is apparent that the accommodation problem for this exhibition is reaching the stage when some alteration in its character or venue, or both, will have to be considered.

The trend of component design continues to maintain the pattern of the last two to three years. In the first place the conventional type of component, which is still being used in extremely large quantities, is continually being improved and made more reliable without necessarily changing its overall characteristics. Thus production is not interfered with too greatly and components remain interchangeable.

New components, however, are continually appearing, mainly for sub-miniature circuit work or for printed circuitry. In the sub-miniature field it must not be thought that all extremely small components are used with transistors; many are used with sub-miniature conventional type valves.

For printed circuitry much attention has been concentrated on ways to produce suitable mounting arrangements for some of the heavier components such as electrolytic capacitors, output transformers, etc. For components such as these, it is unwise to rely on the soldered connections for complete support; hence the need for a "clip-in" or "snap-in" type. Also, a considerable amount of thought has been given to reasonably easy removal of those components with a number of connections of the fairly rigid type. Designs for this type of connection are now appearing.

High temperature operations, miniaturization of transistor components and improved reliability were the three chief trends in the development of components and techniques displayed by the Ministry of Supply. A metal film track made by firing a gold-platinum alloy on to a glass plate results in a very low noise level, high stability and long wearing characteristics. New resistors include a tin-antimony oxide film potentiometer. There was also a 100-watt transistorized power pack developed specially for replacing the usual vibrator power pack. A thousand resistors in pre-determined conditions of temperature and humidity can be tested at a time and the results recorded automatically by a new instrument.

Among the commercial exhibits, an instrument of interest was a universal measuring bridge covering 76 ranges for inductance, capacitance and resistance. It includes a very sensitive valve voltmeter and an internal oscillator for inductance measurements. A mid-scale accuracy of ± 1 per cent. is possible for resistance and capacitance measurements over ranges of 0·1 ohm to 1,000 Mohms and 1 pf to 1,000 μF respectively.

A power amplifier was shown which is designed to meet an increasing demand for a compact drive unit for vibration generators, capable of developing at least 1 kilowatt over a wide frequency range. Its main applications are testing aircraft and electronic components for guided missiles.

The trend towards narrower channel spacing for v.h.f. mobile operation has called for increased stability of the transmitter frequency. A new low voltage crystal oven designed for Style D quartz crystals provides a stability of $\pm 0{\cdot}0005$ per cent. Fundamental and overtone quartz crystals for frequencies between 4 kc/s and 55 Mc/s and, shortly, up to 100 Mc/s, were also exhibited.

New valves and semi-conductor devices included the following:—

Two xenon rectifiers with a new totally enclosed anode and modified anode seal permitting operations at higher gas pressure with increased life and safety margin.

A range of seven new travelling-wave tubes for the 2,000–4,000 Mc/s range. These vary from a 1 mW low noise type (8 db for the 2,000 Mc/s tube and 9·5 db for the 4,000 Mc/s tube) to a 2,000 Mc/s power tube rated at 16 watts.

Low-impedance power oscillator valves for r.f. heating capable of very high outputs at low anode voltages and made especially rugged for factory use.

Two 100-volt sub-miniature diodes intended mainly for industrial applications.

THE PROPERTIES OF SEMI-CONDUCTOR DEVICES*

by

A. A. Shepherd, M.Sc., Ph.D.†

Read before the Institution in London on 24th April, 1957

In the Chair: Mr. G. A. Taylor (Member)

SUMMARY

Rectifiers and transistors developed in recent years make use of rectifying junctions in the body of the semi-conductor, formed by the introduction of appropriate impurities. Junctions may be produced during crystal growth or by alloying or solid state diffusion. The various methods are described for germanium and silicon rectifiers and transistors, and the properties of the resulting devices examined. Semi-conductor photocells are also described.

1. Introduction

Since the invention of the transistor in 1948, research and development in the field of semi-conductors have proceeded at an unprecedented rate. Many new devices of great interest to the circuit engineer have emerged as a result of this large scale effort, the more important of these being rectifiers, transistors and photocells. The technology of the production of these devices involves the preparation of single crystals of germanium and silicon of a high degree of perfection, and the introduction into these crystals of very small amounts of impurities which must be located with great accuracy.

Although early transistors and rectifiers depended on the properties of metal to semi-conductor contacts, in recent years attention has been turned to the use of rectifying junctions in the body of the semi-conductor, formed by the introduction of appropriate impurities. New methods of forming such junctions, such as alloying and, more recently, solid state diffusion, have led to a range of useful circuit elements. In this paper a survey is made of the development of these methods and the properties of the resulting devices.

2. The Properties of Semi-conducting Crystals

2.1. *Intrinsic Semi-conductors*

The two most important semi-conducting materials from the point of view of device fabrication at this time are germanium and silicon. To make satisfactory devices, the material must be prepared in the form of single crystals, which are usually grown from the molten material.

Germanium and silicon are in Group IV of the periodic table, and exhibit a crystal structure similar to that of diamond, that is a tetrahedral structure in which the adjacent atoms are bound by the sharing of the outer valence electrons: each "covalent" bond is formed by the sharing of two electrons. A planar representation of this type of structure, using germanium as an example, is shown in Fig. 1, in which the dots represent the valence electrons of the lattice atoms.

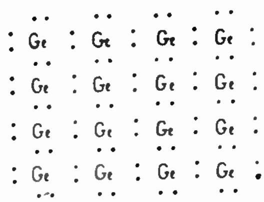


Fig. 1. Planar representation of germanium crystal structure.

Electrical conduction in a pure crystal of germanium can be brought about by the liberation of valence electrons from their normal sites, when they are free to move through the crystal under the influence of an applied electric field. The energy required for such a process is 0·74 eV. This energy may be supplied in a number of ways; the most important from the device aspect is thermal energy and light energy.

* Manuscript received 28th February 1957. (Paper No. 395.)

† Ferranti Ltd., Wythenshawe, Manchester.

U.D.C. No. 621.314.6/.7:537.311.33.

When the electron is removed from the valence level it leaves behind a vacant site which has net positive charge of magnitude equal to the electronic charge. An adjacent valence electron may move into this site, thus effectively transferring the positive site through the lattice. Movement of the vacant site, which is known as a "positive hole," may also give rise to conduction in the material. Thus, in a pure crystal of germanium or silicon, there are two conduction processes, one involving the movement of electrons and the other involving the movement of positive holes. Under these conditions, the carriers are formed in pairs, and the number of electrons is equal to the number of holes. If the density of electrons in the material is n_e and the density of holes is n_h, it may be shown[1] that for thermal formation of electron-hole pairs

$$n_e = n_h = 2\,\frac{(2\pi mkT)^{3/2}}{h^3}\,\exp\left(\frac{E}{2kT}\right) \quad \ldots\ldots(1)$$

where m = electron mass
k = Boltzmann's constant
T = temperature (°K)
h = Planck's constant
E = energy required for the formation of an electron-hole pair.

As the electrons and holes move through the crystal under the influence of an applied electric field, they will be slowed up by the effect of the thermal lattice vibrations. Thus, at a given temperature, there will exist a characteristic value of mobility μ, defined as the mean drift velocity for unit electric field, for electrons and for holes. For pure materials the mobility varies with temperature according to the law

$$\mu \propto T^{-3/2} \quad \ldots\ldots\ldots(2)$$

The contribution of each type of carrier to the electrical conductivity of the semi-conductor is given by

$$\sigma = n\mu \quad \ldots\ldots\ldots(3)$$

where σ is the conductivity, n the carrier density, and μ the carrier mobility. Thus, for a pure or "intrinsic" semi-conductor

$$\sigma = e[n_e\mu_e + n_h\mu_h] \quad \ldots\ldots\ldots(4)$$

where the subscript e refers to electrons and h to positive holes. In practice it is more common to specify resistivity ρ of the material where $\rho = 1/\sigma$; for intrinsic germanium at a temperature of 300°K, the resistivity is 65 ohm-cm while for intrinsic silicon the resistivity at 300°K is 230,000 ohm-cm.

2.2. *Impurity Semi-conductors*

The properties of an intrinsic semi-conductor are characteristic of an absolutely pure substance; the addition of even minute quantities of impurities has a considerable effect upon these properties. For germanium and silicon, the most important impurities are those of Groups III and V of the periodic table.

If an impurity atom of Group V is incorporated into the germanium lattice substitutionally, it will provide an extra valence electron which may be easily removed from the substitutional site; thus every Group V impurity atom acts as a source of one electron which may contribute to the conductivity of the crystal. This case is illustrated in Fig. 2 in which antimony is taken as an example.

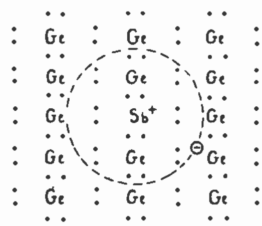


Fig. 2. Substitution of antimony atom in germanium lattice.

Since antimony has five valence electrons, only four of which are needed to satisfy the normal bonds, the extra electron is loosely bound in an orbit which covers several inter-atomic distances; the "activation energy" required to remove the extra electron from this orbit is 0·04 electron volts. The ionization energy of other Group V impurities in germanium is of the same order.

It will be seen that for this case, an electron is made available for conduction without the formation of a positive hole. For germanium or silicon containing a Group V impurity, therefore, the current will be carried mainly by electrons, and the material is designated n-type since negative mobile charge carriers are in the majority. The number of conduction electrons at a temperature T°K in an n-type semi-conductor may be found from the equation

$$n_e = N^{1/2} \frac{(2\pi mkT)^{3/4}}{h^{3/2}} \exp\left(\frac{\Delta E}{2kT}\right) \quad \text{.........(5)}$$

where N is the number of impurity centres per cm³ and ΔE is the appropriate activation energy. Equation (5) is valid only at low temperatures, since as the temperature is raised, more of the impurity centres will become ionized and n_e will approach N. In most material used in devices, the impurity centres will be fully ionized at room temperature. The addition of antimony to germanium in the proportion of 1 in 10^7 is sufficient to reduce the resistivity to about one ohm-cm at room temperature.

When an impurity from Group III of the periodic table is substituted into the lattice of germanium or silicon, only three valence electrons are present at the substitution site and one of the covalent bonds is incomplete. The bond may be completed by the transfer of a valence electron from an adjacent germanium atom, giving rise to a positive hole which may move about the lattice and contribute to the conductivity. The energy required to form the positive hole is of the same order as the ionization energy in the case of the n-type impurity, and the number of holes present at temperature T °K is given by an equation similar to equation (5) where ΔE is now the energy required for the formation of a hole. A semi-conductor in which the majority carriers are positive holes formed in this way is designated p-type.

An important factor in the operation of semi-conductor devices is the behaviour of minority carriers. Under equilibrium conditions, it may be shown that the product of the hole and electron concentrations is given by the relation

$$n_e n_h = 2{\cdot}3 \times 10^{31} \left(\frac{m_e m_h}{m^2}\right)^{3/2} T^3 \exp\left(-\frac{E}{kT}\right) \quad \text{.........(6)}$$

where m_e and m_h are the effective masses of electrons and holes in the crystal and E is the activation energy for production of an electron-hole pair. If the concentration of minority carriers is increased above the equilibrium value by some means, e.g. by illumination, the excess minority carriers will recombine with the majority carriers. For example, if holes are introduced into an n-type semi-conductor, recombination will take place at a rate proportional to the excess hole density. If the excess density is N_h per cm³ then the rate of disappearance of holes will be given by

$$\frac{dN_h}{dt} = -\frac{N_h}{\tau}$$

$$\text{or} \quad N_h = (N_h)_0 \exp(-t/\tau) \quad \text{.........(7)}$$

where $(N_h)_0$ is the initial density of excess holes and τ is a constant which is characteristic of the crystal, known as the minority carrier lifetime. Similar relations will apply for electrons in p-type crystals. Typical values of τ in crystals used for circuit devices range from 10^{-5} to 10^{-3} sec. Lower values of τ are found as the number of crystal imperfections and impurities increase. When there is a spatial gradient of minority carriers these carriers will diffuse in a direction such as to eliminate this gradient. If the characteristic diffusion length for a minority carrier before recombination occurs is L, then

$$L = (D\tau)^{1/2} \quad \text{.........(8)}$$

where D is the diffusion coefficient for the minority carriers in question. In many types of transistor the flow of current across the base region is governed by diffusion of this kind, and the frequency response of the device is dependent on the speed of diffusion.

2.3. *p–n Junctions in Semi-conductors*

If a discontinuity occurs in a single crystal of semi-conducting material so that one part is p-type and the other n-type, the p–n junction forms a potential barrier which has useful device properties. Fig. 3 shows a schematic diagram of such a p–n junction.

If there is initially a uniform concentration of holes in the p-type region and electrons in the n-type region, holes near the junction will flow into the n-type side, together with a similar flow of electrons in the opposite direction, with consequent recombination, until a space charge layer is set up which creates a potential barrier as shown in Fig. 3, preventing the flow of further current.

When the p-type region is made positive with respect to the n-type region, the potential barrier is reduced, and a large current may flow. When the p-type region is made negative, the potential barrier is increased in height.

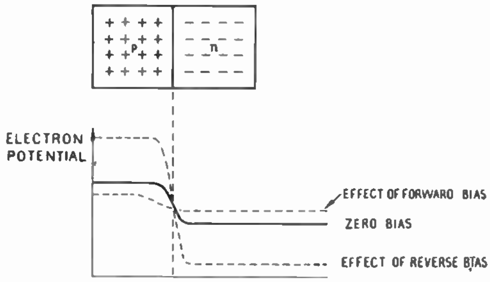


Fig. 3. p–n junction in a semi-conductor showing potential diagram under different conditions of bias.

Under these conditions, the current is carried by thermally generated minority carriers crossing the junction from both sides. Most of the minority carriers produced at a distance greater than a diffusion length from the junction will recombine before reaching it, and thus only those which are generated close to the junction will contribute to the current. This current will be practically independent of applied potential, saturating quickly as the applied potential is increased. The net hole current density i_h across the junction as a function of the voltage V is given by[2]

$$i_h = i_{hs}\left[\exp\left(\frac{eV}{kT}\right) - 1\right] \qquad \text{.........(9)}$$

where V is positive for forward bias, i_{hs} is the reverse saturation current density due to holes and is given by

$$i_{hs} = en_h\sqrt{\frac{D_h}{\tau_h}} \qquad \text{.........(10)}$$

Similar equations hold for electron flow across the barrier. Thus the total current density i will be given by an equation of the form

$$i = i_s\left[\exp\left(\frac{eV}{kT}\right) - 1\right] \qquad \text{.........(11)}$$

where both hole and electron currents contribute to i_s.

The type of current-voltage characteristic predicted by the above considerations is shown in Fig. 4.

The current in the forward direction will increase exponentially with voltage, while the current in the reverse direction will saturate at a value of i_s.

Since i_s is several orders of magnitude lower for silicon than germanium, due to the higher energy required to form electron-hole pairs in silicon, silicon p–n junctions will exhibit much lower reverse saturation currents than germanium junctions, but will also have much lower forward currents at a given voltage.

In most cases of practical interest, the free carrier density on one side of the junction is appreciably greater than that on the other side. For example, if a junction is made with low resistivity p-type germanium and high resistivity n-type germanium, the hole density on the p-type side will greatly exceed the electron density on the n-type side. Under these conditions, the forward current will consist almost entirely of holes passing from p to n while the reverse current will consist almost entirely of holes passing from n to p;

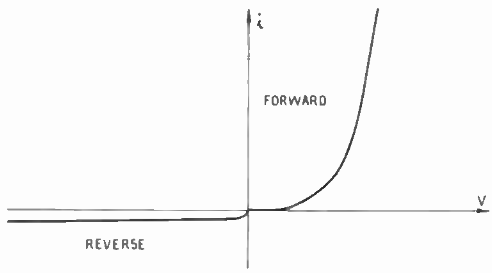


Fig. 4. Current voltage characteristic of p–n junction.
Note.—Change of scale for reverse characteristic.

the number of free electrons on the p-type side is so small that their contribution to reverse current is negligible. Thus a junction of this type when biased in the forward direction forms an efficient emitter of holes into the n-type region. The fate of the injected holes will depend on the minority carrier lifetime in the n-type region. The principle of minority carrier injection at a p–n junction is extremely important in transistors and rectifiers.

When a p–n junction is biased in the reverse direction, the reverse current should saturate at low reverse voltages according to equation

(11). As the reverse voltage applied to the junction is increased, the minority carriers which diffuse into the junction region from both sides will be accelerated across the barrier by the field in the space charge depletion layer. When the field is sufficiently high, these carriers may be able to ionize lattice atoms, producing further carriers which in turn are accelerated to cause more ionization. The process will give rise to a rapid increase in reverse current when a critical voltage is reached, in a manner analogous to the Townsend multiplication process in a gas discharge. The breakdown voltage will depend on the resistivity of the material on both sides of the junction, increasing as the resistivity is made higher.

3. Germanium Junction Rectifiers

3.1. *Grown Junction Rectifiers*

The first type of germanium junction rectifier to be made was the grown junction rectifier, in which a *p–n* junction is produced in the germanium crystal during growth by addition of the appropriate impurities as the crystal is pulled from the melt.

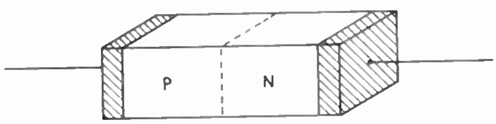


Fig. 5. Construction of grown junction rectifier.

Rectifiers may be made from such a crystal by cutting it into rectangular bars, each containing a *p–n* junction, and attaching end contacts. In this type of device there is an appreciable resistance in series with the junction due to the sections of the bar on each side (see Fig. 5). This has the effect of increasing the forward resistance of the overall device; although the minority carriers injected in the forward biased condition will reduce the effective resistance of the bar in the region of the junction, their mean diffusion length is usually very small (<1 mm), so that the total series resistance is not appreciably affected. Thus, the characteristic of the grown junction type of rectifier has the form shown in Fig. 6. At low forward voltages the limiting resistance is due to the junction itself, but as the forward voltage is increased, the series resistance of the bar limits the current, and the i/V curve becomes a straight line, the slope of which is determined by the series resistance.

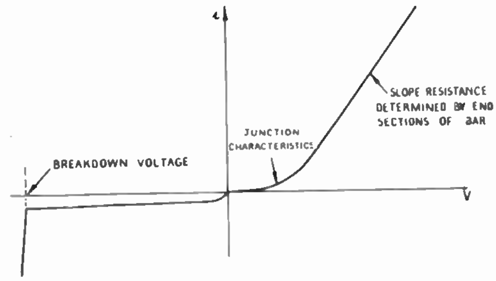


Fig. 6. Characteristic of grown junction rectifier in which the end contacts are separated from the junction by distances much greater than a diffusion length.

It is obviously desirable to reduce the slope resistance as far as possible, and this may be done by

(*a*) reducing the distance between the junction and the end contacts;

(*b*) lowering the resistivity of the material.

The limitation to alternative (*a*) lies in the difficulty of making suitable contacts very close to the junction.

The limitation to alternative (*b*) is that a reduction in the resistivity must be paid for by a decrease in the reverse breakdown voltage, and thus a suitable compromise must be chosen between low forward resistance and high peak inverse voltage.

An additional degree of control over the reverse breakdown voltage in grown junction diodes is afforded by the grading of the junction. An abrupt transition from *p* to *n* results in a lower breakdown voltage than a more graded junction; the grading may be controlled by the rate of growth of the original crystal. The grown junction method is, therefore, suitable for producing rectifiers with high peak inverse voltage ratings, although the forward resistance is high, and the current handling capacity consequently limited.

3.2. *Germanium Alloy Rectifiers*

As a manufacturing method, the grown junction technique has been almost completely superseded by the alloying method for producing rectifiers.[3,4] In this method a button of indium is placed in contact with a thin wafer of n-type single crystal germanium and the temperature is raised to 550°–600°C in an inert atmosphere. The indium button melts and dissolves some of the germanium from the surface of the wafer. After equilibrium is reached, the assembly is slowly cooled and germanium recrystallizes on to the wafer in the same orientation as the parent crystal. The recrystallized material is heavily contaminated with indium and a p–n junction is formed at the interface between the substrate and the regrowth layer, the p-type side of the junction having a very low resistivity, usually in the region of 10^{-2} ohm-cm or lower. A large area base contact is made with an antimony-doped solder on the opposite side of the wafer, and the resulting structure is illustrated in Fig. 7.

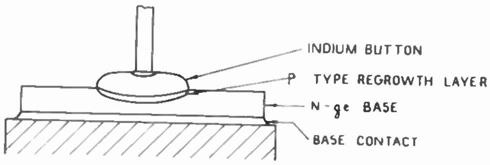


Fig. 7. Schematic diagram of alloy junction germanium rectifier.

This type of rectifier may be constructed so that the distance between the base contact and the p–n junction is small; the series resistance of the base germanium wafer is then small, even when high resistivity material is used. More important, however, is the fact that the thickness of the base layer is comparable with the diffusion length of the holes injected into the base during forward biased operation. These holes, together with the additional electrons drawn in from the base contact to ensure electrical neutrality, may reduce the series resistance of the base at high injection levels to a negligible value. This effect is known as "conductivity modulation" and has the important consequence of permitting the use of high resistivity material for the base region of the rectifier, thus allowing high reverse breakdown voltages to be achieved without the loss of the low forward resistance. This has resulted in the construction of germanium rectifiers capable of withstanding up to 1000 volts in the reverse direction, combined with a forward current density of several hundred amperes per cm² of junction area, with a voltage drop of only half a volt.

It has been shown by Saby[5] that the voltage current characteristic of a modulated rectifier may be described by an equation of the form

$$i = i_s\,[\exp\,(eV'/2kT) - 1] \quad \text{.........(12)}$$

where V' is the voltage across the complete device. It will be seen that this is similar to the relation for the junction itself as given in equation (11), except for the factor of 1/2 in the exponent. At high forward currents departure from this type of characteristic may occur due to self heating.

Alloy junction germanium rectifiers of the form described above are made in many sizes, designed to cover the current range from a few milliamperes to hundreds of amperes per unit. The rectifiers are usually hermetically sealed to protect them from contamination, and for sizes up to tens of amperes mean rectified current, are convection cooled. In the range above one ampere it is usually necessary to employ cooling fins to ensure that the temperature of the rectifying element does not exceed 80°C, which is a practical working limit for germanium. Higher power units employ forced air or water cooling, and water cooled units have been constructed which are capable of handling currents up to 500 amperes.

A typical characteristic for a 100 ampere germanium rectifier is shown in Fig. 8.

Germanium alloy junction rectifiers are beginning to establish themselves in many applications where their small size and weight

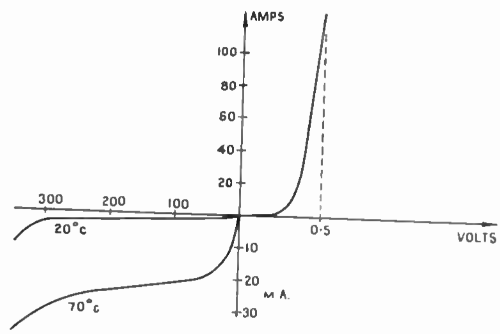


Fig. 8. Typical characteristic of high power germanium rectifier.

are important considerations. The higher power units are already being used in many installations requiring low voltage high current supplies, such as battery charging and electroplating plant. It appears probable that they will also prove of great value in electrical traction applications.

Among the advantages they possess over the conventional selenium and copper oxide rectifiers are their extremely high rectification efficiency due to their low forward voltage drop, indefinite life and absence of any ageing effect on their electrical characteristics.

4. Silicon Junction Rectifiers

4.1. *Grown Junction Rectifiers*

Silicon rectifiers differ from germanium rectifiers in two important respects; they have reverse currents which are several orders of magnitude lower than those for germanium, and forward voltage drop at the junction which is higher than for germanium. These differences are a result of the higher activation energy required to produce electron-hole pairs in silicon.

Referring to equation (11) which gives the voltage-current characteristic of a *p–n* junction, the value of i_s is considerably lower for silicon than germanium. For example, for a silicon rectifier with a very low resistivity *p*-type region and an *n*-type region of 10 ohm-cm resistivity, the reverse current density to be expected, assuming a minority carrier lifetime in the *n*-type region of 10^{-6}sec, is about 5×10^{-9}amps per cm^2 at 300°K. Thus, a grown junction bar 1 mm square would have a reverse saturation current of 5×10^{-11} amps up to the breakdown voltage, as compared with about 10^{-6} amps for a similar germanium junction. In practice, the reverse currents obtained with silicon rectifiers are often two or more orders of magnitude higher than the theoretical value.

As the temperature is increased, the reverse current due to minority carriers crossing the junction (equation (10)) will rise. The product of hole and electron concentrations under equilibrium conditions is given by equation (6); since the majority carrier density effectively saturates at or below room temperature in most devices, it follows that the minority carrier density will increase rapidly with temperature at a rate dependent upon the activation energy E. The activation energy E is 1·12 eV for silicon as compared with 0·74 eV for germanium, and therefore the rate of increase of reverse current with temperature will be higher for silicon junctions. However, the extremely low value of i_s at room temperature means that the junction temperature of the silicon device must be raised to about 250° C. before the reverse current seriously affects the operation of the rectifier. The silicon rectifier therefore, has great advantages at high ambient temperatures. Grown junction silicon rectifiers so far produced have usually been designed specifically for high voltage applications; the graded junctions which may be produced by growing offer possibilities of very high peak inverse voltages, although the forward impedance of this type of rectifier is high. A typical example of a silicon grown junction rectifier, made from a 2 mm square bar, passed a forward current of 50 mA at 6 volts, withstanding a reverse voltage up to 5 kV.

4.2. *Silicon Alloy Rectifiers*

Alloy junction rectifiers have been made in silicon for both low and high power applications. The first account of low power silicon alloy rectifiers was given by Pearson and Sawyer[6]. The method of fabrication described uses an *n*-type silicon base wafer with an aluminium dot or wire alloyed to it to form the rectifying junction. The regrowth layer formed during the alloying process is low resistivity *p*-type due to the high level of of aluminium contamination. The ohmic contact is formed by alloying a gold-antimony electrode to the wafer. The junction area of these early units was about 10^{-3} cm^2; using 20 ohm-cm base material, reverse breakdown voltages up to 300 volts were obtained, with reverse currents of the order of 10^{-5} amps per cm^2 compared with a theoretical value of 10^{-8} amps per cm^2. The units did not exhibit conductivity modulation, the forward currents being about 1 mA at 1 volt.

Later types of aluminium-silicon alloy rectifiers have been described[7] in which a co-axial structure with aluminium and gold antimony wires alloyed to opposite sides of an *n*-type wafer has been used to obtain conductivity modulation. The forward current for this type of structure, with a junction area of 5×10^{-4} cm^2 is about 100 mA at 1 volt. The reverse current is of the same order as for the non-modulated structure, and breakdown voltages in excess of 1000 volts have been achieved by suitable choice of base material. This type of rectifier can handle currents of 100 mA at ambient temperatures up to 75° C, and may be used up to ambient temperatures of 150° C. with suitable derating. The reverse current at this temperature is still only a few microamperes.

A typical characteristic for a 300 volt rectifier is shown in Fig. 9.

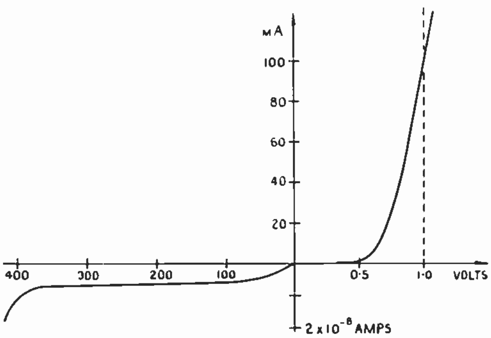


Fig. 9. Typical characteristic for a conductivity modulated silicon alloy rectifier. Junction area 5×10^{-4} cm^2. Temperature 25°C.

The wire alloying method of producing junctions in silicon is not suitable for large area junctions intended for high power rectifiers. The aluminium silicon eutectic formed above the regrowth layer is brittle, and has a different coefficient of expansion from the underlying silicon. Thus if large area junctions are to be made by aluminium alloying, precautions must be taken against mechanical strain arising from this cause. One method which has been used with success is to alloy a thin aluminium foil to an *n*-type base wafer. In this way junctions with a diameter of over $\frac{1}{4}$ in. have been made. An account of the properties of silicon power rectifiers made by this method has been given by Losco[8].

Early units were made from *n*-type wafers having resistivities in the range 1 to 5 ohm-cm. With a junction area of 0·05 cm^2 forward currents up to 10 amps at 1 volt could be obtained. Reverse breakdown voltages up to 100 volts were reported, but the reverse current densities were high, and did not saturate, typical values being about 10^{-2} amps per cm^2 at a reverse voltage of 100 volts.

More recent developments of the large area alloying process aimed at minimizing mechanical strain in the junction region have resulted in a reduction of the reverse currents. Rectifiers capable of working at peak inverse voltages of 300 volts have now been made by this method.

4.3. *Silicon Diffused Junction Rectifiers*

An alternative method for producing large area *p–n* junctions in silicon is by solid state diffusion. In this method, the impurity to be diffused is deposited on the surface of a silicon wafer, which is then heated to a temperature at which the impurity atoms can diffuse through the silicon lattice at a reasonable rate. The concentration of the impurity after diffusion falls off with distance from the surface. If therefore, an *n*-type impurity is diffused into a *p*-type silicon wafer, then at some distance "*x*" from the surface, the concentration of *n*-type diffusant will equal the concentration of *p*-type impurity originally present. A *p–n* junction will thus be formed at a depth "*x*". The diffusion process provides a means of producing planar *p–n* junctions in silicon wafers. The area of the junction is limited only by the size of crystal which may be grown, and the junction is formed without any mechanical strain. The process is, therefore, very attractive for the production of large area rectifiers; since the diffusion process is relatively slow, the position of the junction may be accurately controlled by adjusting time and temperature of diffusion. This in turn means that the uniformity of characteristics in the resulting rectifiers is high.

Silicon diffused junction rectifiers were first reported by Pearson and Fuller[9], and improved types have recently been described by Prince[10]. In these rectifiers the junction is formed by the diffusion of phosphorus into *p*-type base material. The ohmic contact to the opposite side of the *p*-type wafer is made by the diffusion of boron. By using starting material of resistivity greater than 20 ohm-cm breakdown voltages greater than 200 volts may be consistently obtained. Conductivity modulation may be achieved if the wafers are kept thin, to give a base layer thickness of about 0·003 in.

A typical voltage-current characteristic for a diffused silicon rectifier with a wafer 2·5 mm square is shown in Fig. 10. The forward current is 10 amps at 1 volt, and the reverse current increases very little with increasing voltage, the reverse current density being 3×10^{-6} amps per cm² at 25° C, rising to about 10^{-3} amps per cm² at 100° C.

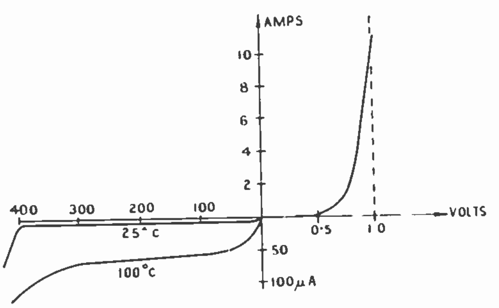


Fig. 10. Characteristics of diffused junction silicon rectifier. Junction area 0·06 cm².

The rating of these rectifiers is determined mainly by the mounting and cooling arrangements, which determine the junction temperature during operation. Providing the junction temperature can be kept below 200° C, the rectifiers can withstand current densities of 1000 amps per cm². A rectifier of the type referred to in Fig. 10 when mounted on a 2½ in. square copper cooling fin, can handle a mean rectified current of 10 amperes at 25° C.

It is to be expected that silicon rectifiers will be used extensively in many types of d.c. power supply. Their very small volume combined with the ability to work up to high ambient temperatures make them very attractive for applications where space and weight saving is of importance, a typical example being aircraft power supply systems. The reverse currents are so low that the power wasted during the reverse half cycle is practically zero, and this fact, combined with low forward voltage drop, give rectification efficiencies above 95 per cent. The uniformity of forward characteristics obtainable with diffused junction rectifiers makes them especially suitable for applications requiring closely matched rectifiers, such as parallel operation to provide very high current output.

4.4. *Silicon Reference Diodes*

The rapid increase of reverse current which occurs due to the onset of breakdown in a *p–n* junction may be used as a voltage reference; the effect is particularly useful in silicon since a very sharp breakdown may be obtained in practical devices and the reverse current before breakdown is negligible for most purposes.

The variation of breakdown voltage with silicon resistivity has been investigated by McKay[11] using alloy diodes with very low *p*-type resistivity. The results obey a law of the form

$$V_B = 80\rho^{0.8}$$

where V_B is the breakdown voltage and ρ is the resistivity of the *n*-type base material used. Thus by a suitable choice of base material it should be possible to fix the reverse breakdown voltage for such a diode. In practice this is possible for low breakdown voltages. It is not yet possible to obtain a high degree of reproducibility at high breakdown voltages, i.e. using high resistivity base material, since surface effects often determine the breakdown characteristics of such diodes.

The slope resistance in the breakdown region is an important parameter of a reference diode, since this governs the degree of regulation obtainable. At very low currents, of the order of a few microamperes, after breakdown, the slope resistance is generally high, but it drops sharply as the current in the breakdown region is increased. It is, therefore, desirable to use reference diodes at a point well into the breakdown region. A typical slope resistance for a small area (5×10^{-4} cm²) alloy diode having a breakdown voltage of 6 volts will be in the region of 10 ohms. As the breakdown voltage is increased by raising the base resistivity, the slope resistance rises. The slope resistance varies approximately as the square of the breakdown voltage. Thus, if a breakdown voltage of 12 volts is required, a lower slope resistance is obtained by using two 6 volt diodes in series rather than a single 12 volt unit. The breakdown voltage varies linearly with temperature over a wide range; for alloy diodes, the coefficient of V_B with temperature varies with the value of V_B, passing through zero at about 6 volts. At lower voltages, the temperature coefficient is negative, i.e. the breakdown voltage falls as the temperature is increased. Also, the breakdown characteristic assumes a more curved shape.

Large area voltage reference units have been made in silicon by the diffusion process, but this type of unit is still in the experimental state. It is to be expected, however, that this type of device will be used for voltage regulation at higher power levels.

5. Junction Transistors

5.1. *Germanium Grown Junction Transistors*

The germanium grown junction transistor was the first form of junction transistor to be produced. As an example of this type of structure, we shall consider first the *p–n–p* configuration shown in Fig. 11. When the emitter *p–n* junction is forward biased, holes are injected into the base region and diffuse across it, being collected at the second *p–n* junction which is reverse biased. In the absence of any current multiplication effects, it follows that the hole current flowing into the collector cannot exceed that flowing out of the emitter. Hence, the current gain in the common base connection cannot exceed unity.

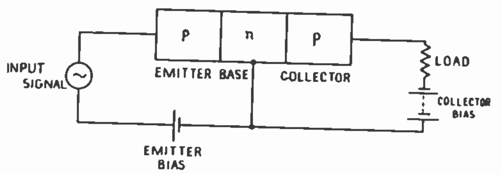


Fig. 11. *p–n–p* grown junction transistor (common base connection).

Since the hole current is introduced in the low impedance emitter circuit, and withdrawn from the high impedance collector circuit, power gain may, however, be obtained. For efficient operation of such a transistor, several important conditions must be fulfilled.

Firstly, it is desirable that as much of the emitter current as possible should consist of holes. This condition may be satisfied by making the *p*-type emitter region of much lower resistivity than the base (see Sect. 2.3). Secondly, it is desirable that as many of the injected holes as possible should reach the collector before recombining with electrons. This may be brought about by making the minority carrier lifetime in the base region as high as possible, by proper crystal growing technique, and by making the base region as thin as possible. Thirdly, the injected holes must be prevented from recombining at the free surfaces of the bar; this is done by careful etching and surface treatment of the junction bar.

For a properly designed and constructed transistor, values of the overall current gain α_{ce} from emitter to collector may be higher than 0·99. If the emitter injection efficiency is high, most of the fraction $1-\alpha_{ce}$ of emitter current which fails to reach the collector will be lost by recombination in the base region or at the surface. The electrons used up in this recombination process are replaced by current from the base lead. Thus, for an emitter current change of ΔI_E, there is a collector current change $\alpha_{ce}\Delta I_E$, and a base current change $(1-\alpha_{ce})\Delta I_E$. If the transistor is connected with the emitter common, and the input signal applied to the base, the current gain obtained is thus $\alpha_{ce}/(1-\alpha_{ce})$. This collector to base current gain is usually designated α_{cb}.

It is important to realise that for this type of transistor when used at low emitter current density, no electric field exists in the base region to assist hole transfer from emitter to collector. The current therefore traverses the base region by diffusion. This fact has important consequences in relation to the frequency response of the device. When a current of holes I_E is injected into the base region from the emitter, a redistribution of charge occurs which is analogous to the process of charging a capacitor. Thus a quantity C_D may be defined, which is the "diffusion capacitance" of the transistor, and is given by

$$C_D = \frac{e}{kT} I_E \frac{W^2}{2D} \qquad \text{.........(13)}$$

where D is the diffusion coefficient for holes in the base region, W is the base thickness and e is the electronic charge.

This diffusion capacitance, combined with the input resistance of the device, may be a limiting factor in determining the frequency response, and must therefore be made as small as possible. From equation (13) it follows that the base thickness W must be reduced for high frequency operation. The collector circuit parameters which limit high frequency response are the collector capacitance, and the series resistance in the collector circuit. The capacitance may be reduced by making the cross-section area of the collector junction as small as possible, and the series resistance by reducing the resistivity of the collector section and making it as short as is practicable. The configuration of a high frequency grown junction transistor is shown in Fig. 12.

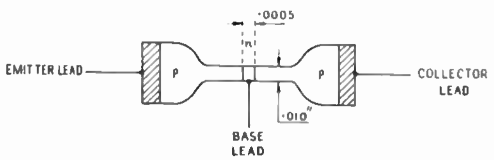


Fig. 12. Configuration of high frequency *p–n–p* grown junction transistor.

It is of importance to note that exactly the same arguments apply to the *n–p–n* type of structure, except that the operating polarities are, of course, reversed. In practice, excellent results have been obtained with both *p–n–p* and *n–p–n* grown junction structures. It is now more common to produce the *n–p–n* type since modern crystal growing techniques make it possible to grow a crystal containing many such junctions.

The criterion usually applied to assess the frequency performance of a transistor is the change in α_{ce} as the frequency is raised. A "cut-off" frequency $f_{c(\alpha)}$ may be defined, at which the value of α_{ce} is reduced by 3 db from the low frequency value, and typical values of $f_{c(\alpha)}$ for germanium grown junction transistors are in the range from 1–10 Mc/s. The width of the base layer to obtain cut-off frequencies in this range must be less than 0·001 in., and under these conditions the main limitation in the frequency response is the series resistance of the very thin base layer, which leads to effective feedback of the collector signal to the emitter circuit.

This limitation has been overcome in a modification of the grown junction structure. A second connection is made to the base layer, to form a tetrode transistor[12]. By applying a few volts of bias between the two base leads, the current crossing the base region from emitter to collector is constricted to flow in a very narrow path close to one of the base leads. By this means, all the transistor action takes place in a channel less than 0·001 in. wide in the immediate vicinity of the operative base lead, so reducing the series resistance in the base region, and preventing appreciable feedback effects. Grown junction tetrodes have been produced on an experimental basis which have values of $f_{c(\alpha)}$ in excess of 100 Mc/s. Such transistors when connected as oscillators will operate at frequencies as high as 1000 Mc/s.

The most suitable field of operation for grown junction triode and tetrode transistors is the high frequency low power field, and several manufacturers have produced transistors for such applications. It is not yet clear whether this type of transistor will eventually be replaced by high frequency transistors made by the alloy or diffusion processes described later in this paper; the final outcome will probably be dictated by the economics of production of the different types.

5.2. *Germanium Alloy Junction Transistors*

In the alloy process for the fabrication of transistor structures, the *p–n* junctions are formed by the recrystallization of germanium from the alloy phase with a suitably chosen metal. The production of *p–n* junctions by alloying was first described by Hall and Dunlap in 1950[13] and has since been used extensively in the preparation of rectifiers (see Sect. 4), and transistors.

The first alloy junction transistors were described in the literature in 1952[14,15], and were of the *p–n–p* type obtained by fusing indium emitter and collector electrodes into an *n*-type germanium base wafer. The configuration of the earlier type of *p–n–p* alloy junction transistor is shown in Fig. 13.

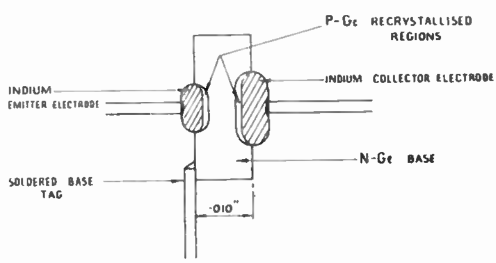


Fig. 13. Sectional view of audio frequency *p–n–p* alloy transistor (germanium).

This type of structure is made by cutting wafers of germanium from an n-type crystal of resistivity about 1 ohm-cm. The wafers are etched and mounted in a jig with two indium buttons positioned as shown in Fig. 13. The assembly is then heated in an inert atmosphere to a temperature in the range 550–600° C. In the heating process, the indium buttons dissolve a little germanium from the wafer which subsequently recrystallizes during the cooling cycle. The recrystallized material is p-type, the resistivity being less than 10^{-2} ohm-cm due to the presence of indium which is incorporated during the recrystallization process. Thus, p–n junctions are formed at the boundaries of the regions which were molten during the alloying cycle.

In practice, the actual junctions may advance a little further due to the diffusion of indium into the n-type base during alloying. The base lead is attached to the wafer by means of antimony doped tin-lead solder.

To produce n–p–n alloy junction transistors in germanium, a similar procedure is used on a p-type base wafer, a lead-antimony alloy being used to produce the emitter and collector junctions[16]. Because p–n–p and n–p–n transistors operate from power sources of opposite polarity, the two types may be advantageously used together in special circuits[17].

Alloy junction transistors differ from the grown junction types described in Section 5.1 in several important respects. The p–n junctions are discontinuous or "step" junctions as opposed to the more graded type of junction obtained in the grown junction transistor. This is an advantage in the case of the emitter, since it leads to a very high emitter efficiency, especially since the emitter region is of very low resistivity. The step collector junction, however, has a higher capacitance per unit area than the grown junction type of collector, which limits the operating frequency, and the peak collector voltage for a given base resistivity is lower than that for a grown collector. The series resistance in the emitter and collector circuits is much smaller in the alloy transistor than in the grown transistor, since direct connection is made to the emitter and collector through the alloyed indium buttons, whose resistance is negligible.

Typical characteristics for the p–n–p alloy junction type of transistor described above are:—

Low frequency α_{ce}	·98
α cut-off frequency	100 kc/s
Power dissipation	100 mW
Max. collector bias	30 V.

Although these transistors are suitable for audio frequency operation, it is necessary to take special precautions in construction if good high frequency operation is to be achieved. Recent work aimed at producing high frequency alloy transistors has been concerned with

(*a*) reduction of the base resistance by suitable constructional geometry,

(*b*) formation of closed spaced, planar alloy junctions.

One such investigation has been described by Herold[18]. In the high frequency p–n–p transistor described by Herold, the base resistance is reduced by using a low resistivity base wafer (0·7 ohm-cm); the base wafer is made 0·020 in. thick, and the collector is alloyed in a well formed in the base layer. In this way, the series resistance of the base is kept low, while the effective part of the base may be made very thin. By the use of this type of construction, the base resistance may be reduced by a factor of about 5 below that of a conventional alloy transistor, and the diffusion capacitance (see Sect. 5.1) by about 20. Values of the cut-off frequency of 5–10 Mc/s are typical.

In another method for reducing the base resistance of the transistor, a ring base contact of gold-antimony alloy is fused to a thin base wafer around the collector electrode. This method requires careful control of the alloying process to achieve satisfactory wetting of the germanium in an annular region by the ring base. By careful attention to the alloying process it has recently been possible to produce alloy junctions up to 0·030 in. in diameter and flat to within 0·0001 in. Junction spacings of less than 0·001 in. have been obtained, giving cut-off frequency in excess of 4 Mc/s[19].

The alloy transistors so far described are suitable for low power applications where the collector dissipation does not exceed a hundred milliwatts. Under these conditions, the injected carrier density does not greatly exceed

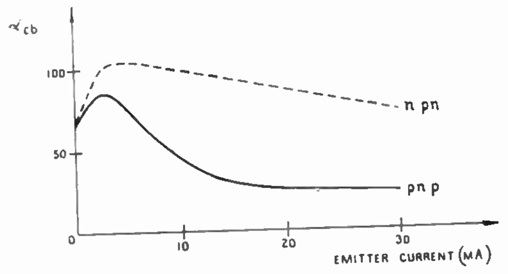


Fig. 14. Variation in α with emitter current.

the base impurity concentration. In transistors designed for higher power use, the same constructional principles apply, but the high current density has an appreciable effect on the current gain. If we use as an example, the *p–n–p* structure, the injected holes from the emitter require an equal number of electrons to maintain charge neutrality in the base layer. Thus, at high current densities, there is an electron concentration in the base which may exceed the number of Group V impurities present. The effect is the same as a reduction in base resistivity, and an appreciable fraction of the emitter current is now carried by electrons, which reduces the effective value of α. The extra electrons in the base layer will in some cases increase the recombination rate for the injected holes, leading to a further reduction in α. However, the presence of the extra electrons in the base layer results in a field gradient at high current densities which aids hole flow towards the collector. This effect tends to maintain a high value of α; the overall result of these competing factors is shown for a *p–n–p* structure in Fig. 14, in which the variation of α_{cb} with emitter current is plotted. For *n–p–n* transistors the reduction of α at high emitter currents is smaller, as may be seen from the dotted curve of Fig. 14.

In transistors designed for high power use, high values of α can thus be obtained by increasing the cross section area of the electrodes so that the current density is kept reasonably low. A practical limit is set by the operating frequency required.

For medium power applications, larger versions of the alloy junction transistors described above have been used with success[20]; one of the main problems involved is the removal of several watts of power dissipated in the transistor without allowing the temperature of the device to exceed the operating limit, which in most cases is in the region of 70° C. Removal of the dissipated power has been achieved by the use of copper heat sinks in thermal contact with the transistor, and by the use of liquid filled encapsulations. For high power applications structures have been used in which the emitter and collector are in the form of intermeshing grids on one side of the germanium wafer, while the base contact is made to the opposite side[21]. Transistors of this type have been successfully operated at output levels in excess of 100 watts.

5.3. *Germanium Surface Barrier Transistors*

The surface barrier transistor depends for its action on the properties of the contact between a metal and a perfectly clean germanium surface. This type of transistor has been specifically designed for low power, high frequency applications[22].

If metal electrodes are deposited on a clean germanium surface, they act in the same way as *p–n* junctions, i.e. a metal electrode in intimate contact with *n*-type germanium acts as an efficient emitter or collector electrode. In the surface barrier transistor, a wafer of *n*-type germanium is electro-etched by means of narrow jets of electrolyte, so that the finished wafer has two concentric depressions, one on each side, separated by a thin web which forms the base layer of the transistor. Emitter and collector electrodes are then electroplated on to the bottom of the depressions. The electro-etching process is a very accurate method of shaping the germanium, allowing base thicknesses in the region of 0·0002 in. to be formed in a reproducible manner. Typical dimensions for a surface barrier transistor are shown in Fig. 15.

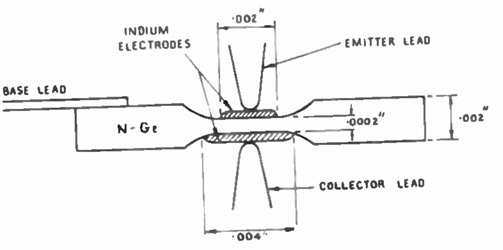


Fig. 15. Typical dimensions of surface barrier transistor.

The surface barrier transistor has a very good high frequency performance, and values of α cut off frequency in excess of 50 Mc/s are common. However, due to the fact that the power dissipation takes place in an extremely thin section of the wafer, the power handling capacity of these transistors is very limited, a typical value being 10 mW.

5.4. *Germanium p–n–i–p and n–p–i–n Transistors*

The *p–n–i–p* transistor and its counterpart, the *n–p–i–n* transistor, were designed at the Bell Telephone Laboratories for very high frequency applications[23]. In the straightforward *p–n–p* transistor, as has already been stated, one of the major limitations in high frequency operation is the fact that the emitted holes cross the base region by a diffusion process, i.e. there is virtually no electric field to assist their transfer from emitter to collector, and the movement is therefore, slow. If the base region is made very thin to shorten the transit time, it fails to support the collector voltage and an effective short circuit, known as "punch through" occurs from collector to emitter.

In the *p–n–i–p* transistor, the problem is solved by the insertion of an intrinsic layer of semiconductor between a very thin base layer and the collector. After the injected holes have crossed the base layer, they are swept across the intrinsic layer, since a high field exists in this region due to the high resistivity. Thus the effective transit time is determined by diffusion through a very thin base region, and the collector voltage is supported by the intrinsic layer. In a typical *p–n–i–p* transistor, an *n*-type low resistivity base layer about 0·0001 in. thick is grown on a high resistivity "intrinsic" section of crystal (usually slightly *n*-type) which is cut to a thickness of 0·0007 in. Indium emitter and collector electrodes are applied to opposite sides of the resulting wafer by jet electroplating and alloying to produce the configuration shown in Fig. 16. A ring base electrode is attached by alloying to the thin base layer.

p–n–i–p and *n–p–i–n* transistors are capable of very high frequency operation, α cut-off frequencies of up to 200 Mc/s having been measured.

Because of the rather thicker wafers which may be used, higher power levels are possible

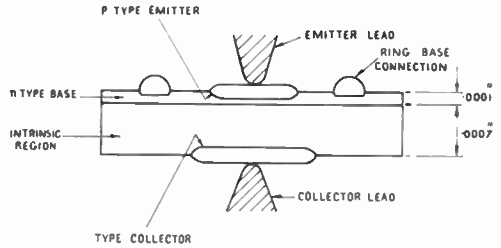


Fig. 16. *p–n–i–p* transistor structure.

than for the surface barrier type of transistor. For example, up to 100 mW of output power may be obtained at 100 Mc/s. It seems rather unlikely that the *p–n–i–p* transistor will be suitable for large scale production; recent advances in diffusion technology make it probable that for very high frequency applications, the drift transistor produced by diffusion will be the preferred type for mass production.

5.5. *The Germanium Drift Transistor*

In this type of transistor, as with the *p–n–i–p* transistor, very high frequency operation is achieved by a sweeping field which assists the transfer of carriers from emitter to collector. In the case of the drift transistor however, this is brought about by a grading of the base resistivity, the resistivity being low at the emitter side and high at the collector side. In a recently announced[24] type of drift transistor produced by the diffusion technique, an *n*-type impurity is diffused into a *p*-type wafer. The *p–n* junction so formed is used as the collector junction, and the *n*-type diffused layer has a resistivity gradient which assists hole transfer towards this junction. A *p*-type electrode is alloyed on to the *n*-type surface to form an emitter, together with a suitable *n*-type electrode to form the base contact. The area round these electrodes is then masked and the diffused layer is etched away from the rest of the wafer to reduce the area of the collector

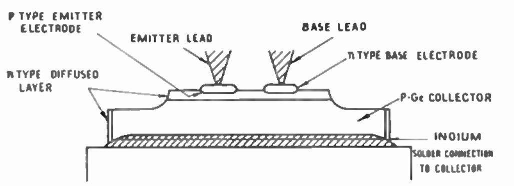


Fig. 17. Diffused type of drift transistor.

junction. The resulting configuration is shown in Fig. 17. The emitter and base electrodes measure 0·001 in. × 0·006 in. and the base thickness is about 0·00006 in.

Performance figures reported for these transistors quote low frequency values of α_{ce} in excess of 0·98, with an α cut-off frequency greater than 500 Mc/s. Power handling capacity in excess of 100 mW should be obtainable with this structure without exceeding the limiting junction temperature, and it is forecast that in the near future, operation at frequencies up to 1000 Mc/s will be possible.

5.6. *Unipolar or Field Effect Transistors*

The unipolar or field effect transistor operates on a principle different from that for any of the transistors previously described; the name unipolar is used since carriers of one sign only are involved in the transistor action[25]. It consists of a semi-conductor bar whose resistance may be varied by altering its effective cross sectional area. This is done by constricting the channel through which current carriers may flow by means of electrodes on the sides of the bar. Fig. 18 shows a schematic diagram of a unipolar transistor. An n-type germanium bar has two low resistivity p-type sections introduced opposite to one another as shown. If the p-type sections are negatively biased with respect to the n-type bar, a space charge depletion layer extends from the junctions into the interior of the bar, its thickness depending on the bias voltage. The space charge depletion regions on each side of the central channel restrict the width of this channel and thus, the current flowing through it may be modulated by the voltage on the "gate."

Because of the potential drop in the bar due to the electron current flowing from source to drain, the effective bias on the drain end of the gate is higher than that on the source end, and the channel is tapered, being narrower at the drain end. If a high enough bias voltage is applied to the gate, the channel is pinched off, and the drain current saturates. The resulting characteristics are shown in Fig. 19, in which the drain current is plotted as a function of drain voltage for various values of gate bias V_G. The voltages are measured with respect to the source end of the bar.

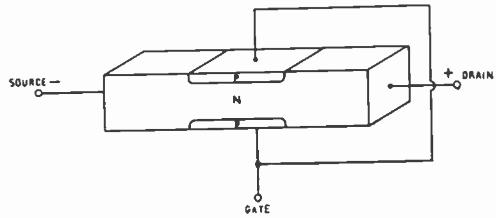


Fig. 18. Structure of unipolar transistor.

Since little power is consumed in the input circuit because the gate is reverse biased, the device has a high input impedance; the output impedance is also fairly high. The frequency response obtainable is higher than for a conventional transistor structure of similar dimensions, and these transistors have good power handling capacity at high frequency, typical values being 300 mW for a transistor having a frequency cut-off of 100 Mc/s.

In the fabrication process, the junctions may be formed by the alloying or diffusion techniques previously described. At the present time these transistors have not reached the production stage, although considerable experience has been gained on a laboratory scale.

5.7. *Silicon Transistors*

Since silicon transistors will operate at a higher junction temperature than germanium transistors, they offer the possibility of operation at higher ambient temperatures or alternatively, at higher dissipation limits than similar germanium devices. The grown junction silicon transistor is the only type which is established in production at this time; the alloy type silicon transistor has given rise to many problems in technology which do not occur with germanium, and is only now beginning to reach large scale production in the U.S.A.

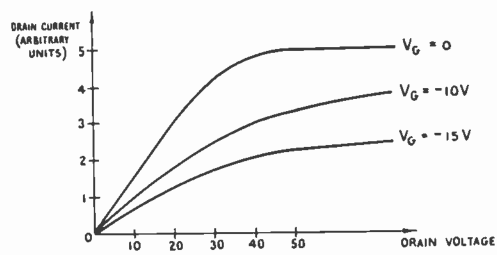


Fig. 19. Characteristics of unipolar transistor.

The general principles of fabrication for a silicon grown junction transistor are similar to those outlined for germanium in Section 5.1. The "rate growing" process, in which several *n–p–n* structures may be produced in one crystal is widely used; silicon grown junction transistors produced by this method have been described which are capable of working at power dissipation of 1·5 watts, with an α cut-off frequency of 5 Mc/s[26].

Although the desirability of the silicon alloy junction transistor has been long evident, the fabrication problems are much more difficult than for germanium transistors. In making the alloy junctions, an alloying agent must be used which does not introduce strains near the junction due to expansion mismatch; it must also be capable of wetting the silicon easily during the alloying cycle. In addition to these requirements, the alloying cycle should not cause appreciable deterioration of the minority carrier lifetime in the silicon. Another major difficulty experienced in making silicon transistors is the prevention of carrier recombination at the free surfaces of the base layer; this problem is more serious in silicon than in germanium, and leads to a low value of α. In a recently developed *n–p–n* audio frequency transistor described by Nelson[27], techniques have been developed which have surmounted these difficulties, and audio frequency transistors have been made which have values of α_{cb} at low frequency of up to 30, and low frequency power gain of 40 db. Operating temperatures of up to 150° C. are possible. Audio frequency silicon transistors of *n–p–n* and *p–n–p* types are, however only in limited production at the present time.

Low power, high frequency silicon transistors have been made by a modification of the surface barrier technique[28] in which a silicon wafer is shaped by electro-etching, and subsequently has aluminium emitter and collector electrodes evaporated on and alloyed by a "micro-alloying" technique which enables base layer thicknesses after alloying of about 0·0005 in. to be produced. Using this construction, transistors have been made with an α cut-off frequency of up to 12 Mc/s, and a low frequency α_{cb} of about 12 (higher values have been obtained in transistors not specifically designed for switching). Operating temperatures up to 150° C. are allowable for these transistors.

The technique of solid state diffusion has recently been applied to the production of high frequency silicon transistors[29]. In these transistors, both the base layer and the emitter are formed by diffusion into a silicon wafer, which forms the collector electrode. This is achieved by the simultaneous diffusion of aluminium and antimony into an *n*-type wafer. The aluminium diffuses faster than the antimony, but has a lower surface concentration. Thus, the result is an *n–p–n* configuration which forms the basis of the transistor. Connection is made to the emitter layer by electrically bonding a gold-antimony plated tungsten point to the surface, and connection to the base is made by alloying an evaporated aluminium electrode through the *n*-type surface layer into the underlying *p*-type base region. This base connection does not short the base to the emitter, since the *p*-type regrowth layer formed on alloying the aluminium is rectifying to the emitter region.

The structure is masked and etched to reduce the collector area as for the germanium drift transistor already described.

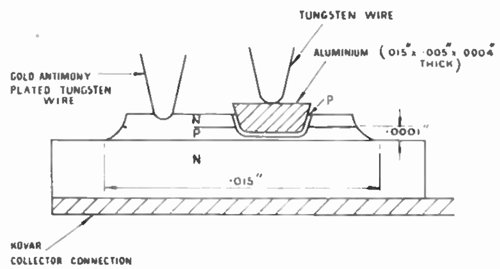


Fig. 20. Construction of diffused silicon transistor.

Transistors of this type, shown in Fig. 20, have an α cut-off frequency in excess of 100 Mc/s, a low frequency α_{ce} of 0·97, can handle up to 400 mW of power, and, suitably derated, may be used at temperatures up to 150° C. The graded base layer produced by diffusion is an important factor in improving the frequency response. Although these transistors have as yet only been produced on a laboratory scale, they show considerable promise, especially if the technique of making the base layer connection can be simplified.

6. Semi-conductor Photocells

Most semi-conductor photocells in current use employ a body of semi-conducting material containing a p–n junction. These cells fall into two main classes, namely photoconductive and photovoltaic types. In the photoconductive cell, a p–n junction biased in the reverse direction has its characteristic modulated by incident light. In the photovoltaic cell, the built in field at the p–n junction is used to separate the holes and electrons produced by incident light.

6.1. *Photoconductive Cells*

The grown junction photoconductive cell consists of a bar of semi-conductor with a p–n junction at right angles to its length, provided with ohmic contacts at the ends[30, 31]. When such a cell is biased in the reverse direction, a saturation reverse current will flow in the absence of illumination which consists of holes crossing the junction from the n-side and electrons from the p-side. The current density due to holes is given by $en_h\sqrt{(D_h/\tau_h)}$ where n_h is the density of holes in the n-type region and D_h is the hole diffusion coefficient. τ_h is the hole lifetime in the n-type region near the junction. The reverse current density due to electrons is given by a similar expression, giving a total reverse saturation current density

$$i_s = e\left[n_h\sqrt{\left(\frac{D_h}{\tau_h}\right)} + n_e\sqrt{\left(\frac{D_e}{\tau_e}\right)}\right]$$

If the p–n junction is illuminated, the production of electron hole pairs by photons gives rise to increased values of n_e and n_h, thus increasing the saturation current. For the carriers to be effective, they must be produced within a diffusion length L of the p–n junction, given by

$$L = (D\tau)^{\frac{1}{2}}$$

It follows that for optimum photo response, a high value of minority carrier lifetime is desirable on both sides of the junction.

A typical configuration for a grown junction photocell is shown in Fig. 21: a bar of dimensions 0·04″ × 0·04″ × 0·125″ is cut from a germanium crystal containing a p–n junction. The junction is centrally disposed as shown.

After attaching ohmic contacts to the ends of the bar, the unit is etched, and encapsulated in a plastic block with one clear face to allow the passage of light to the element. Using tungsten light of colour temperature 2800°K, the sensitivity of such a photoconductive cell is about 30 microamperes per millilumen. For reverse bias voltages above 1 volt, the photocurrent saturates well and the increase in photocurrent with light flux is linear over a wide range. This linearity is useful where high fidelity reproduction of light signals is required. Typical voltage-current characteristics are shown in Fig. 22.

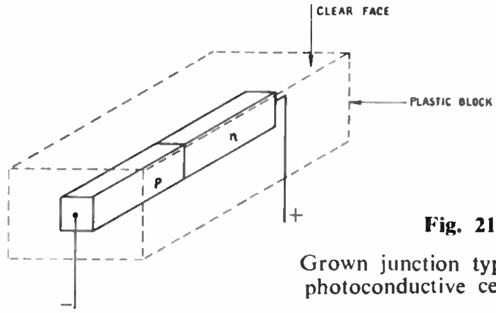


Fig. 21. Grown junction type photoconductive cell.

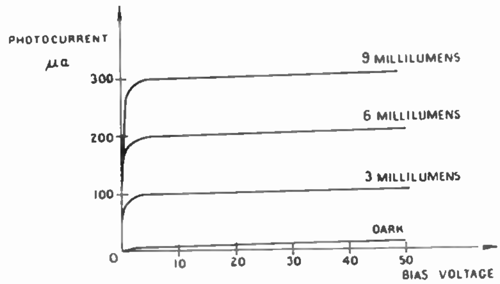


Fig. 22. Typical voltage current characteristics for a grown junction photoconductive cell.

Since the photocurrent is due to the collection of minority carriers produced near the junction, it will be seen that the sensitivity of the cell will fall off as the illuminated area is moved away from the junction; if the sensitivity is S when the junction itself is illuminated, it will fall to S/e at a distance L from the junction, where L is the minority carrier diffusion length on the side of the junction in question.

Since the carriers move to the junction by a diffusion process, the frequency response of the cell will be determined by the diffusion time. Thus, for high frequency response, the illuminated area should be as narrow as possible and centred on the junction itself. If the junction is masked to give an illuminated width of 0·010 in., response times of less than one microsecond can be achieved.

The spectral sensitivity of the cell is usually limited at the ultra violet end by the encapsulating plastic, which ceases to transmit in the ultra-violet; at the infra-red end of the spectrum, the sensitivity falls when the long wavelength absorption limit of the germanium is passed, at about 11,000 angstroms.

It has been found by Goucher[32] that for light of wavelength 1000–2000 angstroms, electron-hole pair formation in germanium takes place with a quantum yield very close to unity. Therefore it cannot be expected that the sensitivity of a single junction photocell will be improved by increasing the efficiency of the activation process. Increased sensitivity may be obtained in a multi-junction cell designed so that multiplication of the primary photocurrent is obtained.

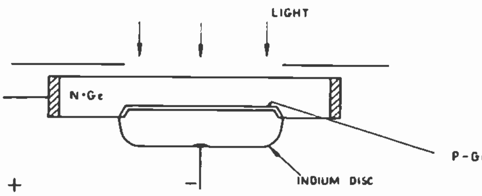


Fig. 23. Alloy junction germanium photocell.

Another form of junction photocell is the alloy junction cell in which the junction is produced by alloying an indium disc into a thin wafer of germanium. The construction of such a cell is illustrated in Fig. 23. In this type of cell the holes produced as a result of illumination of the upper surface of the germanium wafer diffuse to the junction, and are collected there to produce a current in the external circuit. Since the entire surface illuminated is at the same distance from the junction, the positioning of the light to obtain maximum sensitivity is not so critical as for the grown junction cell previously described. For maximum sensitivity it is essential that the thickness of the wafer should not exceed the mean diffusion length for the holes. A thin germanium section is also necessary if high frequency response is to be achieved, since the frequency response is governed by the same considerations outlined in Section 5.2 for the alloy junction transistor.

The thin germanium wafer gives rise to a fall-off in the long wavelength sensitivity, since the layer is too thin for the complete absorption of long wavelength radiation, some of which is consequently lost as a source of photocurrent. A typical value for the sensitivity of such a cell is 30 microamperes per millilumen, measured with a tungsten source of colour temperature 2800°K.

6.2. *Photovoltaic Cells*

In the photovoltaic cell, there is no external applied voltage as for the photoconductive cell. The irradiation of the junction with light produces a photovoltage across the device which can be used to drive current through an external circuit. If a *p–n* junction is irradiated with light under open circuit conditions, the built-in voltage at the junction separates the electrons and holes which are produced by the incident light. The holes are drawn to the *p*-type side of the junction, and the electrons to the *n*-type side. This separation of electron-hole pairs causes a voltage to appear across the device in the direction of forward bias of the *p–n* junction. For most germanium cells this is about 0·5 volts, while for silicon it can be somewhat higher.

Photovoltaic cells are usually used to produce a current in some external circuit, and since the current is proportional to the light collecting power of the device, it is desirable that as large a junction area as possible should be exposed to the incident radiation. One of the most promising applications of the semiconductor photovoltaic cell is in the conversion of solar energy to electrical power. It has been shown[33] that for a silicon junction photocell, a limiting efficiency of about 22 per cent. may be expected in converting solar to electrical energy. Silicon approaches the optimum material for this purpose, and silicon solar cells have been made by creating a *p–n* junction in a wafer of *n*-type material by boron diffusion. The junction is placed about 0·0001 in. from the surface and acts as an efficient means of separating the electron hole pairs formed by the sun's radiation; with cells of this kind conversion efficiency of over 11 per cent. has been achieved, and this represents by far the most efficient method of conversion of solar energy to electrical power. The output voltage of such a cell on load is about 0·3 volts under conditions of maximum power transfer, and silicon solar cells are ideal as a charging supply for telephone system storage batteries in remote locations. They have already been in use in trial installations in the U.S.A.

7. Conclusion

The development of semi-conductor device technology over the past few years has led to a wide range of rectifiers, transistors and photocells. Many of these devices have emerged as a result of the evolution of new methods of production of *p–n* junctions in semi-conductor crystals; the increasing measure of control over the position of such junctions has given rise to new standards of high frequency performance and uniformity of device characteristics. The advent of solid state diffusion as a method of device fabrication may be expected to have far reaching consequences in these aspects of semi-conductor device technology.

Although many of the devices discussed in this paper are in the preliminary stages of development, it is to be expected that during the next few years they will become available on a large scale to the circuit engineer, and place at his disposal the many advantages which they offer.

8. References

1. R. H. Fowler, "Statistical Mechanics," 2nd ed. (Cambridge University Press, 1936.)
2. W. Shockley, "Electrons and Holes in Semi-conductors, with applications to Transistor Electronics. (Van Nostrand, New York, 1950.)
3. R. N. Hall, "Power rectifiers and transistors," *Proc. Inst. Radio Engrs*, **40**, pp. 1512-1518, Nov. 1952.
4. T. H. Kinman, G. A. Carrick, R. G. Hibberd and A. J. Blundell, "Germanium and silicon power rectifiers." *Proc. Instn Elect. Engrs*, **103**, Part A, pp. 89-107, April 1956.
5. J. Saby, "Junction Rectifier Theory," Physical Society Spring Meeting, April 1956. (Published in "Report of Semi-conductor Conference," 1957.)
6. G. L. Pearson and B. Sawyer, "Silicon *p–n* junction alloy diodes," *Proc. Inst. Radio Engrs*, **40**, pp. 1348-1351, Nov. 1952.
7. D. E. Mason, A. A. Shepherd and W. M. Walbank, "Silicon junction power diodes," *J.Brit.I.R.E.*, **16**, pp. 431-441, August 1956.
8. E. F. Losco, "Silicon power rectifier handles 1,200 watts," *Electronics*, **27**, pp. 157-159, Dec. 1954.
9. G. L. Pearson and C. S. Fuller, "Silicon *p–n* junction power rectifiers and lightning protectors," *Proc. Inst. Radio Engrs*, **42**, p. 760, April 1954.
10. M. B. Prince, "Diffused *p–n* junction silicon rectifiers," *Bell Syst. Tech. J.*, **35**, pp. 661-684, May 1956.
11. K. G. McKay, "Avalanche breakdown in silicon," *Phys. Rev.*, **94**, pp. 877-884, 15th May 1954.
12. R. L. Wallace, Jr., L. G. Schimpf and E. Dickten, "A junction-transistor tetrode for high-frequency use," *Proc. Inst. Radio Engrs*, **40**, pp. 1395-1400, Nov. 1952.
13. R. N. Hall and W. C. Dunlap, "*p–n* junctions prepared by impurity diffusion," *Phys. Rev.*, **80**, pp. 467-468, 1st Nov., 1950.
14. R. R. Law, C. W. Mueller, J. I. Pankove (Pantchechnikoff) and L. D. Armstrong, "A developmental germanium *p–n–p* junction transistor," *Proc. Inst. Radio Engrs*, **40**, pp. 1352-1357, Nov. 1952.
15. J. S. Saby, "Fused-impurity *p–n–p* junction transitors," *Proc. Inst. Radio Engrs*, **40**, pp. 1358-1360, Nov. 1952.
16. D. A. Jenny, "A germanium *n–p–n* alloy junction transistor," *Proc. Inst. Radio Engrs*, **41**, pp. 1728-1734, Dec. 1953.
17. G. C. Sziklai, "Symmetrical properties of transistors and their applications," *Proc. Inst. Radio Engrs*, **41**, pp. 717-774, June 1953.
18. E. W. Herold, "New advances in the junction transistor," *Brit. J. Appl. Phys.*, **5**, pp. 115-126, April 1954.
19. J. J. Ebers, "Alloyed-junction transistor development," *Bell. Lab. Rec.*, **34**, pp. 8-12, Jan. 1956.
20. B. N. Slade, "Transistors I," p. 153. (R.C.A. Laboratories, Princetown, 1956.)
21. R. N. Hall, "Power rectifiers and transistors," *Proc. Inst. Radio Engrs*, **40**, pp. 1512-1518, Nov. 1952.
22. "The Surface-barrier Transistor." A series of 5 papers by Philco Research Division. *Proc. Inst. Radio Engrs*, **41**, pp. 1702-1720, Dec. 1953.
23. J. M. Early, "*p–n–i–p* and *n–p–i–n* junction transistor triodes," *Bell Syst. Tech. J.*, **33**, pp. 517-533, May 1954.
24. C. A. Lee, "A high-frequency diffused-base germanium transistor," *Bell Syst. Tech. J.*, **35**, pp. 23-34, Jan. 1956.
25. W. Shockley, "A unipolar 'field-effect' transistor," *Proc. Inst. Radio Engrs*, **40**, pp. 1365-1376, Nov. 1952.
26. M. Tanenbaum, L. B. Valdes, E. Buehler, and N. B. Hannay, "Silicon *n–p n* grown-junction transistors," *J. Appl. Phys.*, **26**, pp. 686-692, June 1955.
27. H. Nelson, "Transistors I," p. 172 (R.C.A. Laboratories, Princetown, 1956).
28. A. D. Rittmann and T. J. Miles, "High-frequency silicon alloy transistors," *Trans. Inst. Radio Engrs (Electron Devices)* **ED-3**, No. 2, p. 78, April 1956.
30. J. N. Shive, "Properties of the M-1740 *p–n* junction photocell," *Proc. Inst. Radio Engrs*, **40**, pp. 1410-1413, Nov. 1952.
31. J. N. Shive, "The properties of germanium phototransistors," *J. Opt. Soc. America*, **43**, pp. 239-244, April 1953.
32. F. S. Goucher, "The photon yield of electron-hole pairs in germanium" (Letter), *Phys. Rev.*, **78**, p. 171, July 1950.
33. J. J. Loferski, "Theoretical considerations governing the choice of the optimum semi-conductor for photovoltaic solar energy conversion," *J. Appl. Phys.*, **27**, pp. 777-784, July 1956.

APPLICANTS FOR ELECTION AND TRANSFER

As a result of its May meeting the Membership Committee recommended the following elections and transfers to the Council.

In accordance with a resolution of Council and in the absence of any objections, the election or transfer of the candidates to the class indicated will be confirmed fourteen days after the date of circulation of this list. Any objections or communications concerning these elections should be addressed to the General Secretary for submission to the Council.

Direct Election to Member

GAVIN, Professor Malcolm Ross, M.B.E., M.A., D.Sc. *Bangor.*

Transfer from Associate Member to Member

WEECH, Lt.-Col. Charles William Thacker, R. Sigs. *Baghdad.*

Direct Election to Associate Member

ATHANASSIADES, Lt.-Col. Elias. Royal Greek Army. *Athens.*
SCHROEDER, Lieut. Henry Alexander, R.N. *Elgin.*
SINCLAIR, Sqdn. Ldr. Douglas Alan, R.A.F. *Longley Lane.*
SLATER, Frank. *St. Albans.*

Transfer from Associate to Associate Member

DOYLE, Comdt. Eamon Daly, Eireann Army. *Clondalkin.*
TITMUSS, Walter Ernest. *Surbiton.*

Transfer from Graduate to Associate Member

BARNES, George Raymond. *Ainsworth.*
CASS, Wilfred Richard. *Cambridge.*
LENT, Stuart James. *Tadworth.*
MITAL, Roop Narayan, B.Sc., B.E.(Hons.). *Calcutta.*
SIMMONS, Brian Desmond. *London, S.E.10.*
SINHA, Flt. Lt. Kailash Nath, B.Sc., I.A.F. *Poona.*
TUCK, Leslie. *Johannesburg.*
WALDRON, Richard Arthur, B.A.(Cantab.). *Chelmsford.*
WILKINSON, Sqdn. Ldr. Donald Edgar, R.A.F. *Waddington.*

Direct Election to Associate

DWYER, William John. *Hounslow.*
HEATHCOTE, Duncan. *Salisbury, Southern Rhodesia.*
HUMPOLETZ, Justus Ernest, B.Sc. *Harrow.*
PEARSON, Philip Macdonald. *Manchester.*

Transfer from Student to Associate

LICHTBLAU-PORGES, Adolf. *Sutton, Surrey.*
PABRAI, Om Prakash. *Bombay.*
THORP, Philip David. *Bletchley.*

Direct Election to Graduate

CHINN, Henry Ronald. *Bentleigh, Victoria.*
ROACH, Peter Francis. *Chelmsford.*
WILLIAMS, Brian Edward Frank. *London, S.W.16.*

Transfer from Associate to Graduate

BOICE, Cyril James. *Anglesey.*

Transfer from Student to Graduate

DARVELL, John Louis. *London, N.13.*
GRAY, Robert Frank. *Weybridge.*
THOMAS, Edison Symonds, B.Sc.(Hons.). *Bangalore.*

STUDENTSHIP REGISTRATIONS

AYO, John Ernest. *Harrow Weald.*
BROWN, John Charles Mutch. *Aberdeen.*
BURROW, George. *Ulverston.*
CURLEY, Michael Joseph Colum. *Ballaghaderreen, Eire.*
DONENFOLD, Adolph. *Kinat Motzkin, Israel.*
GARRIOCH, James Traill Moodie. *Sanday, Orkney.*
GEORGE, Paul Maliakal, B.Sc. *Jullundur City.*
GOODCHILD, John Lionel. *Shoeburyness.*
GURSOY, Ali Mustafa. *Nicosia.*
HOPKINS, Roland Michael Terrence. *Headington.*
HULATT, John. *Wyton, Hunts.*
LIM, Teck Siong. *Singapore.*
MATIUR-RAHMAN, Plt. Off. A. K. M., P.A.F. *Sargodha, West Pakistan.*
MEDIWAKE, Weerakoon Bandara. *Colombo.*
MUSCAT, Aurelio. *Libya.*
PLANT, Victor Ernest. *London, E.17.**
PREM MOHAN. *Roorkee.*
RAMAKRISHNA SASTRY, K. V., M.Sc. *Madanapalle.*
RICHARDS, James Dudley, M.Sc. *Whitley Bay.*
SANKHALA, Ranjeet Singh. *Veraval, Bombay.*
SHARMA, Jagdish Chander, B.A. *Delhi.*
SHARMA, Kamlaker, B.Sc. *Bareilly.*
SHARMA, Sh am Sunder. *Kaithal.*
SHUKLA, Trilok Nath. *Rohtak.*
SINGH, Sukdev, B.Sc. *Delhi.*
SKINNER, Dennis Grant. *Ilford.*
SWARUP, Pande Bishun, B.Sc. *New Delhi.*
SWEMMER, Benjamin Northling. *Johannesburg.*
THORNTON, Edward Alan. *Bristol.*
TIN, Da id Wa- Cheong. *Kowloon, Hong Kong.*

* Reinstatement.

THE DESIGN OF PHASE-LINEAR INTERMEDIATE-FREQUENCY AMPLIFIERS*

by

A. van Weel, Dr. Techn. Sc.†

SUMMARY

Data are given for the design of i.f. amplifiers with flat group-delay characteristics for combinations of from one to six tuned circuits. Both staggered-tuned single circuits and combinations of band filters are considered. Figures of merit are calculated and compared with the figures for networks based on flat-amplitude design.

1. Introduction

The design of intermediate-frequency amplifiers for a flat amplitude characteristic, based on the combination of a number of staggered-tuned single circuits or coupled circuits ("band filters") or both can be found in many publications and textbooks on wide-band amplifiers. The shape of the phase characteristic is usually not considered, although it is well known that the distortion of the transient response of such an amplifier, caused by phase errors, can be much larger than the circuit distortion due to limited irregularities of the amplitude characteristic.

Some papers have been published on phase-linear i.f.-amplifier circuits.[1,2,3] However, the results are usually not given in a form in which they can be readily applied to the practical design of i.f. amplifiers. The present paper aims at filling this gap in the literature.‡

We shall calculate the group delay time ($d\varphi/d\omega$) characteristics of the transfer impedances rather than the phase-angle characteristics, the reason being that we are finally interested in the phase deviations of the modulation of the transmitted signal and not in the phase shift of the carrier frequency. These phase deviations can directly be calculated from the group-delay characteristics.[4]

Cascades of various numbers of single tuned circuits will be considered; the possibility of replacing two symmetrically staggered-tuned single circuits by a band-filter will be taken into account. Figures of merit for the gain-bandwidth product are calculated.

Sections 2-6 give the calculations and considerations underlying the design of flat group-delay amplifiers containing from one to six tuned circuits. The data given in these sections are discussed in Sections 7 and 8. The optimum design for different numbers of stages is treated in Section 9; the last section (10) mentions some practical aspects of the tuning procedure of phase-linear amplifiers.

All calculations are based on a small relative bandwidth.

2. One Tuned Circuit

The impedance of a circuit consisting of an inductance L, a capacitance C and a resistance R_0 in parallel is given by

$$\mathbf{Z}_0 = \frac{1}{1/R_0 + j\omega C + 1/j\omega L} = \frac{R_0}{1 + jx}$$

where x, the frequency-dependent co-ordinate, is given by

$$x = \omega_0 R_0 C\left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}\right) \cong 2\Delta\omega R_0 C.$$

In cascades with several tuned circuits, the index n will indicate circuits with different amounts of detuning. The index 0 will always denote the most detuned circuit, and n will increase with less detuned circuits. The frequency-dependent co-ordinate x will always

* Manuscript received 15th March 1957. (Paper No. 397.)

† Philips Research Laboratories, N.V.Philips' Gloeilampenfabrieken, Eindhoven, Netherlands.

U.D.C. No. 621.397.621.54:621.375.121.1.

‡ Note added in proof.—The reader is also referred to the following article which appeared after the completion of this paper: Yona Peless and T. Murakami, "Analysis and synthesis of transitional Butterworth-Thomson filters and bandpass amplifiers," *R.C.A. Rev.*, **18**, pp. 60-94, March 1957.

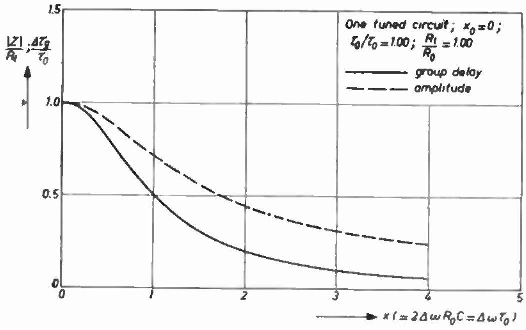


Fig. 1.

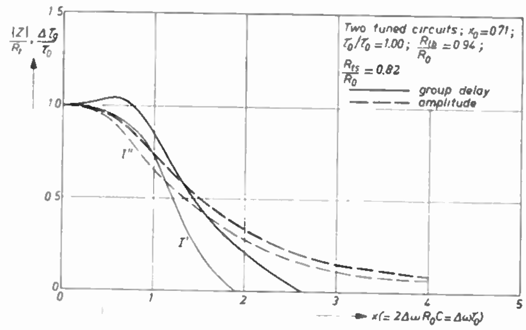


Fig. 2.

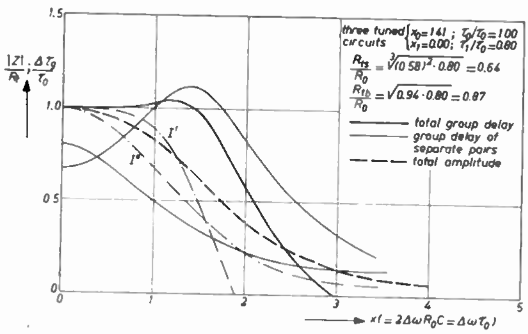


Fig. 3.

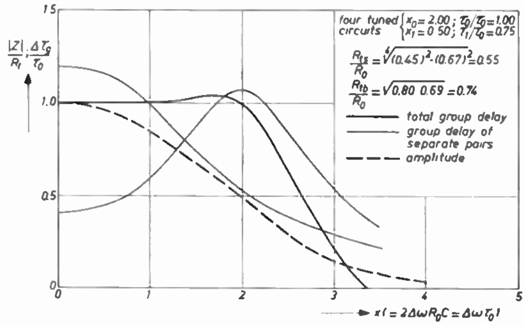


Fig. 4.

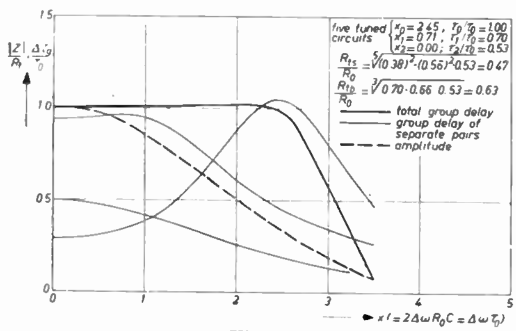


Fig. 5.

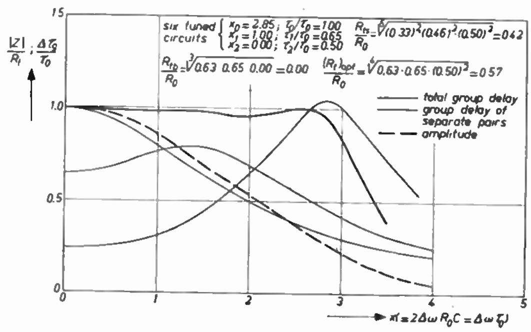


Fig. 6.

Figs. 1 6. Group-delay and amplitude characteristics for combinations of 1–6 tuned circuits, designed for a flat group-delay characteristic. The thin lines give the group-delay curves for the separate circuit pairs. The curves I′ and II′ in Figs. 2 and 3 hold for a maximally flat design.

The insets give the design data (x_n, τ_n/τ_0) and the effective transfer impedance at centre frequency for staggered-tuned stages (R_{ts}/R_0) and for combinations of band filters (R_{tb}/R_0) respectively.

be expressed in R_0, as defined by the last equation. In later sections, the parameter x_n will be used; this parameter is defined by $x_n = 2\Delta\omega_n R_n C$. The detuning is always given with respect to the centre frequency of the pass-band. The phase angle of the impedance of a single circuit is given by

$$\tan(-\varphi) = x$$

and therefore the group-delay follows from

$$\tau_g = \frac{d(-\varphi)}{d\omega} = \frac{d(-\varphi)}{dx} \cdot \frac{dx}{d\omega} = \frac{2R_0C}{1+x^2} = \frac{\tau_0}{1+x^2}$$

Here, τ_0 gives the maximum value of the group-delay, which occurs at the resonance frequency.

Group-delay and amplitude characteristics for a single tuned circuit are given in Fig. 1.

3. Two Tuned Circuits

The transfer impedance of two staggered-tuned circuits is determined by

$$\mathbf{Z}_0' \cdot \mathbf{Z}_0'' = \frac{R_0}{1+j(x-x_0)} \cdot \frac{R_0}{1+j(x+x_0)} = \frac{R_0^2}{1+x_0^2-x^2+2jx}$$

where $x_0 = 2\Delta\omega_0 R_0 C$ is a measure of the amount of detuning of each circuit as compared with the central frequency. With a band-filter, x_0 equals the product of the coefficient of coupling and the quality of the circuit ($x_0 = k_0 Q_0$).

The phase angle and group-delay follow from

$$\tan(-\varphi) = \frac{2x}{1+x_0^2-x^2}$$

$$\tau_g = \frac{2\tau_0}{1+x_0^2} \cdot \frac{1+\dfrac{x^2}{1+x_0^2}}{1+2\dfrac{1-x_0^2}{1+x_0^2}\cdot\dfrac{x^2}{1+x_0^2}+\dfrac{x^4}{(1+x_0^2)^2}}$$

with again $\tau_0 = 2R_0C$.

The "maximally-flat" situation in a design for flat amplitude characteristics is often defined as the situation where the maximum number of derivatives of $|\mathbf{Z}|$ by x is zero for $x=0$. This leads to the situation of "critical detuning" or, with a band-filter, of "critical coupling," which occurs for $x_0 = 1$.

The same definition for the group-delay curve yields the value

$$x_0^2 = 1/3.$$

A number of amplitude and group-delay characteristics for various values of x_0 are depicted in Fig. 7. Normalized ordinate and abcissa scales have been used; the magnitude of the mean effective transfer impedance per stage is given by

$$R_t = \frac{R_0}{\sqrt{(1+x_0^2)}}$$

In the case of a "maximally-flat" amplitude curve, $x_0^2=1$, a rather large variation in the value of the group-delay time can be seen to occur within the "amplitude frequency band," which extends in this case roughly from $x=-1$ to $x=+1$. An increase of the "amplitude bandwidth" can be brought about by choosing a larger value for x_0; the amplitude curve with $x_0^2=2$ is reasonably flat to $x=+1{\cdot}6$. However, the shape of the group-delay curve has deteriorated and the maximum group-delay variation within this frequency band is twice as large as was the case with $x_0^2=1$.

With a "maximally-flat" group-delay curve, $x_0^2=1/3$, the amplitude curve can be seen to vary a good deal within the band of constant group-delay (in this case up to $x=\pm 0{\cdot}6$). Here too, an increase of the "group-delay bandwidth" can be brought about by an increase of x_0; with $x_0^2=1/2$, the group-delay curve is flat up to $x=\pm 0{\cdot}9$. This time, however, the shape of the "other" curve, the amplitude characteristic, has improved, although not to the same extent as the group-delay curve.

We thus conclude that an increase of the coupling or detuning above the value that is critical for a flat amplitude characteristic, increases the amplitude bandwidth but deteriorates the group-delay curve, whereas an increase of the coupling or detuning above the value that is critical for a flat group-delay characteristic increases the group-delay bandwidth and improves the amplitude curve. It is obviously advantageous to aim at this last overcritical adjustment.

The amplitude and group-delay characteristics for a case with the slightly over-critical value of $x_0^2=1/2$ (the critical value of $x_0^2=1/3$) are given in Fig. 2. Since we are interested in the absolute variation of τ_g/τ_0 and not in the relative variation as is the case with the amplitudes, the scale of $\Delta\tau_g/\tau_0$ along the vertical axis has been placed so as to indicate $\Delta\tau_g/\tau_0=1$ for $x=0$ in all figures. This does therefore not imply that $\tau_g=\tau_0$ for $x=0$. In Fig. 2, τ_g is larger than τ_0 for $x=0$, as follows directly from a comparison with Fig. 7, curve $x_0^2=1/2$. The same holds for Figs. 3-6.

The curves I′ and I″ in Fig. 2 give group-delay and amplitude curves for the case of maximally-flat group-delay.

4. Three Tuned Circuits

The transfer impedance follows in this case from

$$\mathbf{Z}_0'.\mathbf{Z}_0''.\mathbf{Z}_1 = \frac{R_0^2\ R_1}{(1+x_0^2-x^2+2\mathrm{j}x)\,(1+\mathrm{j}\,\frac{\tau_1}{\tau_0}\,x)}$$

because the third circuit is for reasons of symmetry tuned to the centre frequency. The group-delay is given by

$$\tau_g = \frac{2\tau_0}{1+x_0^2}\cdot\frac{1+\dfrac{x^2}{1+x_0^2}}{1+2\dfrac{1-x_0^2}{1+x_0^2}\cdot\dfrac{x^2}{1+x_0^2}+\dfrac{x^4}{(1+x_0^2)^2}} + \frac{\tau_1}{1+(\frac{\tau_1}{\tau_0})^2\,x^2} = \left(\frac{2\tau_0}{1+x_0^2}+\tau_1\right)\cdot\frac{1+A\,x^2+B\,x^4}{1+C\,x^2+D\,x^4+E\,x^6}$$

with

$$A = \frac{1+(1-x_0^2)\,\tau_1/\tau_0+(1+x_0^2)\,(\tau_1/\tau_0)^2}{1+\frac{1}{2}(1+x_0^2)\,\tau_1/\tau_0} \qquad B = \frac{\frac{1}{2}\,\tau_1/\tau_0+(\tau_1/\tau_0)^2}{(1+x_0^2)\{1+\frac{1}{2}(1+x_0^2)\,\tau_1/\tau_0\}}$$

$$C = \frac{2\,(1-x_0^2)}{(1+x_0^2)^2}+\left(\frac{\tau_1}{\tau_0}\right)^2 \qquad D = \frac{1+2\,(1-x_0^2)\,(\tau_1/\tau_0)^2}{(1+x_0^2)^2} \qquad E = \frac{(\tau_1/\tau_0)^2}{(1+x_0^2)^2}$$

The design for a maximally-flat group-delay can be found by making the coefficient A equal to C and B equal to D. This yields

$$x_0^2 = 0{\cdot}95$$

$$\frac{\tau_1}{\tau_0} = 0{\cdot}79$$

The corresponding group-delay and amplitude characteristics are shown in Fig. 3, curves I′ and I″.

Better properties for both characteristics can be expected with an over-critical adjustment, in which the group-delay curve is allowed to oscillate between certain limits. The design of flat-amplitude networks in similar circumstances is often based on a Tchebyscheff approximation of the ideal flat characteristic. Such an approximation meets with some difficulties in the case of a group-delay curve, because the group delay is given by a quotient of two polynomials, of which the number of coefficients is greater than the number of parameters in the circuit.

A much shorter way to a useful design in this situation is a trial-and-error method. For this purpose, the author tried out a number of combinations, taking in each case one of the shouldered group-delay curves of Fig. 7 and combining it with the most suitable single-circuit curve. In this way he obtained the amplitude and group-delay curves given by the heavy lines in Fig. 3 as being optimum. Fig. 3 also gives the group-delay curves of the doublet and the single circuit each separately.

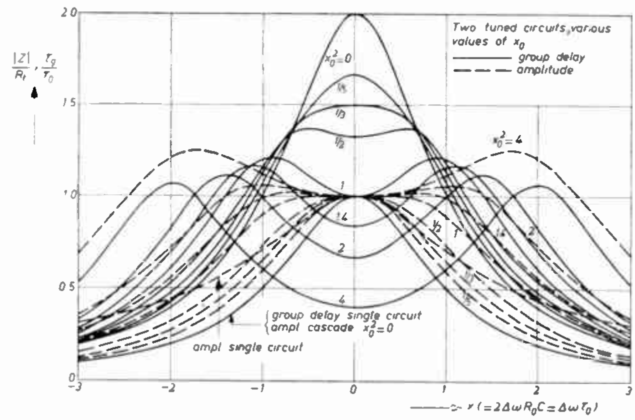


Fig. 7. Group-delay and amplitude characteristic of a band filter or a combination of two staggered-tuned circuits for various values of the coupling or detuning parameter $x_0=k_0Q_0$ or $x_0=Q_0\,.\,2\Delta\omega_0/\omega$ respectively. The amplitude curves are normalized with respect to the transfer impedance at centre frequency R_t.

5. Four Tuned Circuits

The theoretical approach follows the same lines as in the case of three tuned circuits. However, the calculation of the design data for the maximally-flat group-delay curve is rather cumbersome. The author refrained from carrying through this tedious computation because the optimum design data for practical use will in any case differ again from those for a maximally-flat group-delay characteristic. To find a usable over-critical design, a number of combinations of one of the shouldered over-critical group-delay curves of Fig. 7 with a less-than-critical curve were tried out; in this way the characteristics given in Fig. 4 were found.

6. Five and Six Tuned Circuits

Both theoretical and trial methods become increasingly impractical with a larger number of tuned circuits. However, the experience gathered with the four preceding designs, together with the marking of certain systematical changes in the data with increasing numbers of circuits made it possible to find designs which in any case give an increased group-delay bandwidth (Figs. 5 and 6). No guarantee, however, can be given for these designs being optimum.

7. Discussion of the Data Given in Figs. 1—6

The normalized co-ordinate scales in Figs. 1 - 6 permit the calculation of the maximum group-delay variations, if the amplitude bandwidth is given: the values of $\Delta\tau_g/\tau_0$ can be read from the figures, τ_0 is determined from $\Delta\omega_{max}\,\tau_0 = x_{max}$ and x_{max} is again given by the figures and the fall of the amplitude curve for which the amplitude bandwidth is defined.

The absolute value of the group-delay is usually not of any importance; it can be calculated from

$$\tau_g = \sum_0^n \frac{\tau_n}{1+x_n^2}$$

The relative amplitude variations within and outside the pass-band can be read from the figures. The absolute value of the mean transfer impedance at centre frequency does not directly follow from the figures, because the ordinate scale is normalized relative to R_t, which is the mean transfer impedance per stage at centre frequency, whereas the abscissa scale is normalized relative to R_0, which is the resonance impedance of the most detuned circuits. The insets of Figs 1-6 give the value of R_t/R_0; however, R_t for the case of staggered-tuned circuits is different from R_t for a band filter. Therefore the values of R_t/R_0 for these two different situations have been given separately by R_{ts}/R_0 and R_{tb}/R_0 respectively. The same total capacitance per stage has been assumed in both cases; with a band filter, this capacitance has been presumed to be evenly divided over primary and secondary circuit.

To give information on the transfer impedances of the separate stages, the value of R_t/R_0 is in Figs. 3-6 first given in the shape of an nth-power root (n being the total number of stages) of a product of factors, each factor standing for the value of R_t/R_0 for one of the stages. The first factor of the product holds for the most detuned circuits, the second factor for the most-but-one detuned circuits, etc.

The numerical values for x_n and τ_n/τ_0 given in the insets contain all other information necessary for the design, as will be illustrated with an example in Section 9.

8. Comparison of Flat-Amplitude and Flat-Group-delay Designs

The group-delay and amplitude curves of Figs. 1-6 are given together in Fig. 8. In the same way, Fig. 9 gives both kinds of characteristics in the case of a flat amplitude design. These figures demonstrate that it is not possible to design a network containing minimum-phase types of circuit only with both flat amplitude and flat group-delay characteristics. However, the object of an amplifier is in general to pass the signal without distortion in amplitude or in phase, and although the emphasis has in the past certainly been laid too much on the amplitude characteristic only, it is of course at least quite as wrong to base the design on the group-delay curve only and to neglect the shape of the amplitude characteristic.

For an optimum design, we have to aim at the situation where the total influence of amplitude and group-delay distortion on the transient response is a minimum; it can be expected that this will be the case if both influences are equal. However, the conditions for such an optimum do depend on the type of modulation used to transmit the signal proper. We will consider only amplitude modulation; even then the optimum design will be different for a double-sideband modulated signal or for a vestigial-sideband modulated signal.

In order to permit a comparison of the gain for the optimum designs which we are going to discuss with the more familiar flat-amplitude designs, we shall calculate figures of merit for both cases. The figure of merit is given by

$$F = 2\Delta\omega_{max} R_t C = x_{max} \cdot \frac{R_t}{R_o}$$

To evaluate this figure of merit, we have to define the maximum allowable amplitude and group-delay distortion in the case of a flat-amplitude design and a flat-group-delay design respectively; these maximum distortions should have influences of the same magnitude on the transient response. The frequency limits for these two cases will be indicated by $x_{max\,a}$ and $x_{max\,g}$ respectively.

We will discuss the choice of these maximum deviations for double-sideband and vestigial-sideband modulated signals separately.

8.1. *Double-Sideband Modulated Signal*

With a double-sideband modulated signal, the width of the pass band is often defined as the bandwidth between the frequencies where the amplitude curve has fallen a factor of $\sqrt{2}$. According to the principle that relative amplitude variations and absolute phase variations are comparable as regards the magnitude of their influence on the transient response, the maximum allowable deviation from a linear phase characteristic should be 0·3 radian.

Up to now we have only considered the group-delay characteristics; in the ideal situation the group-delay curve will be flat throughout the whole pass-band, in which case the phase-frequency characteristic is linear. The phase-angle deviations from a linear characteristic can for each sideband component be found from a given group-delay curve provided the carrier frequency is known. With a double-sideband modulated signal, the carrier frequency coincides with the centre frequency of the pass band, and the phase-angle deviation for a side-band component can be calculated from

$$-\Delta\varphi = \int_{\omega_c}^{\omega} \left\{ (\Delta\tau_g)_{\omega} - (\Delta\tau_g)_{\omega c} \right\} d\omega$$

$$= \int_{\omega_c}^{\omega} \left\{ \left(\frac{\Delta\tau_g}{\tau_0}\right)_{\omega} - \left(\frac{\Delta\tau_g}{\tau_0}\right)_{\omega c} \right\} d\omega$$

$$\tau_0 = \int_{x_c}^{x} \left\{ \left(\frac{\Delta\tau_g}{\tau_0}\right)_{\omega} - \left(\frac{\Delta\tau_g}{\tau_0}\right)_{\omega c} \right\} dx$$

where ω_c or x_c and ω or x denote the carrier frequency and sideband frequency respectively.

We performed this integration graphically for all group-delay characteristics in Figs. 8 and 9 by measuring the area between the $\Delta\tau_g/\tau_0 = 1$ line and each group-delay curve at various values of x. The results are given by the chain-dotted lines in Figs. 8 and 9. The values of $x_{max\,a}$ (amplitude $\sqrt{2}$-times smaller) and $x_{max\,g}$ (phase-angle deviation 0·3 radian) can now be read in these figures; these values are tabulated in Tables 1 and 2.

Table 1

Values of $x_{max\,a}$ and $x_{max\,g}$ with a flat group-delay design

Number of circuits	1	2	3	4	5	6
$x_{max\,a}$	1·00	1·06	1·32	1·42	1·46	1·50
$x_{max\,g}$	1·17	1·70	2·38	2·92	3·34	3·58

Table 2

Values of $x_{max\,a}$ and $x_{max\,g}$ with a flat amplitude design

Number of circuits	1	2	3	4	5
$x_{max\,a}$	1·00	1·41	2·00	2·63	3·23
$x_{max\,g}$	(1·17)	(1·60)	1·80	2·18	2·52

It can be concluded from these tables, that for a double-sideband modulated signal the optimum design with equal amplitude and phase distortion is somewhere between a flat amplitude design and a flat group-delay design; it is nearest to the maximally-flat amplitude design. It should be taken into account that the flat amplitude designs of Fig. 9 are maximally-flat networks and not over-critically adjusted designs.

8.2. *Vestigial-Sideband Modulated Signal*

In amplifiers for a vestigial-sideband modulated signal the situation is quite different. The carrier frequency has to be laid at a point where the amplitude characteristic has fallen by a factor of two, i.e. on the flank of the intermediate frequency amplitude characteristic. This characteristic is therefore not flat in this region, whereas the overall amplitude characteristic for the modulation frequencies can and should be flat. Obviously no phase errors can be allowed in this frequency band, particularly because the sideband frequencies in this part of the pass-band correspond to low modulation frequencies where phase errors are most objectionable.

There is no possibility in this case of deriving the magnitude of allowable phase errors from the shape of the amplitude characteristic, and we are therefore obliged to make a more or less arbitrary choice. To this purpose we define a phase-angle deviation of 0·01 radian as acceptable; this can be compared to an amplitude distortion of 1 per cent. We could use the phase angle curves of Fig. 8 to find the corresponding values of $x_{\max g}$. However, the fact that all group-delay curves in Fig. 8 have roughly the same shape will enable a direct determination of $x_{\max g}$ from these group-delay curves as can be shown as follows.

The upper part of all group-delay curves in Fig. 8 can roughly be approximated by a straight line under 45 deg. passing through the point where the group-delay curve has fallen to the value of 0·9. This is exemplified with curve 6 in Fig. 8. The area of the hatched triangle equals the phase-angle deviation; the magnitude of this area is $\frac{1}{2} . 0{\cdot}2 . 0{\cdot}1 = 0{\cdot}01$ radian, i.e. the maximum tolerated phase deviation. From this it follows that the group-delay bandwidth is limited at the points where the group-delay curve crosses the line $\Delta\tau_g/\tau_0 = 0{\cdot}9$.

We mentioned already that the choice of $\Delta\varphi = 0{\cdot}01$ radian is rather arbitrary. From the foregoing it will be clear that this choice corresponds to the point where the phase-angle characteristic begins to show substantial deviations from linearity. An increase of the tolerable phase deviation yields but small increase in bandwidth as can be judged from the fact that for curve 6 values of $\Delta\varphi = 0{\cdot}01$ and $0{\cdot}04$ correspond to values of $x_{\max g} = 2{\cdot}94$ and $3{\cdot}05$ respectively. We consider therefore the choice of a maximum phase-angle deviation of 0·01 radian as a good workable proposition.

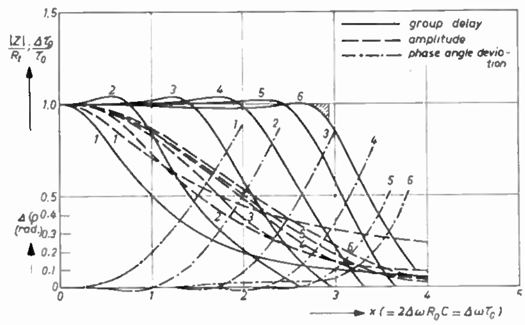


Fig. 8. Group-delay and amplitude characteristics for combinations of 1–6 tuned circuits, designed for at flat group-delay curve. The chain-dotted lines give phase-angle deviations from a linear phase-frequency characteristic.

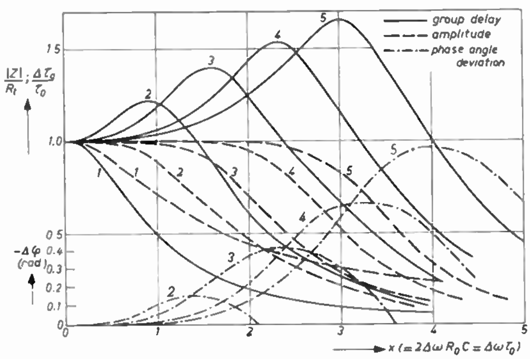


Fig. 9. Group-delay and amplitude characteristics for combinations of 1–5 tuned circuits, designed for a flat amplitude curve. The chain-dotted lines give phase-angle deviations from a linear phase-frequency characteristic.

In order to compare the group-delay bandwidth according to this definition with the bandwidth that is acceptable from amplitude point of view, we have to define a maximum allowable drop in amplitude. This can only be a guess, because now the situation at the carrier-frequency side of the pass-band is quite different from the situation at the other side, which corresponds to the highest modulation frequencies. For the moment we arbitrarily maintain the earlier definition of $x_{\max a}$ for the frequencies where the amplitude has fallen to $1/\sqrt{2}$ of the maximum value.

The values of $x_{\max}$ can be read from Figs. 8 and 9 and the figures of merit can be calculated. The results are given in Tables 3 to 5. For simplicity's sake, the values of $x_{\max g}$ for the flat amplitude designs of Fig. 9 were taken at the points where the group-delay curve shows the same deviation of $0{\cdot}1.(\Delta\tau_g/\tau_0)$. The group-delay curves in these designs (Fig. 9) have a different shape as compared with those of Fig. 8 and the phase-angle deviation will therefore be different (in fact somewhat larger due to the more gradual increase of $\Delta\tau_g$).

Table 3 holds for a flat group-delay design and different numbers of stages, giving the figures both for staggered-tuned circuits and for band filters. Table 4 holds for a flat-amplitude design for staggered-tuned circuits only. It should be taken into account that this table holds for a maximally-flat design; a more over-critical design could yield more gain. Table 5 gives figures for cascades with different numbers of stages, using in every stage the same slightly over-critically coupled band filter ($x_0^2 = 1{\cdot}4$). The overall bandwidth for the total cascade is again defined by the 3 db points.

The combination of different band filters for a flat-amplitude design is not considered, because such a design is very rare in practice, a circumstance which might be due to the large phase errors which are present in such an amplifier. However, the gain per stage for this kind of amplifier will certainly be larger than in any of the other cases.

In all three tables, the total capacitance per stage has been assumed to be the same; with band filters this capacitance is again presumed to be evenly divided over input and output terminals.

The following conclusions can be drawn from these tables:

(1) In the flat group-delay design (Table 3), the bandwidth is for one and two tuned circuits more limited by the group-delay curve than by the amplitude curve, $x_{\max g}$ being smaller than $x_{\max a}$. For combinations of more circuits, the amplitude curve limits the bandwidth.

In the flat amplitude design the limitation by group-delay deviations is always more stringent than the amplitude restriction.

(2) The figure of merit for flat group delay, F_g, is in the design for flat group delay, Table 1, always better than in the two other designs, Tables 4 and 5. This was of course to be expected.

(3) The value F_a for flat amplitude is largest for the cascade of band filters of Table 5. The values of F_a are roughly the same for the staggered-tuned amplifier, based on amplitude (Table 4) as for the combination of band filters designed for flat group-delay (Table 3. The smallest values of F_a are found in the staggered-tuned amplifier, designed for flat group-delay.

Table 3

Figures of merit for combinations of 1 to 6 tuned circuits, designed for a flat group-delay characteristic. The value of $x_{\max g}$ holds for $\Delta\tau_g/\tau_0 = 0{\cdot}1$; the values of $x_{\max a}$ hold for a 3 db drop in amplitude.

Number of circuits	R_t/R_0		$x_{\max g}$	$F_g = x_{\max g} \,.\, R_t/R_0$		$x_{\max a}$	$F_a = x_{\max a} \,.\, R_t/R_0$	
	Staggered tuned	Band filters		Staggered tuned	Band filters		Staggered tuned	Band filters
1	1·00	—	0·35	0·35	—	1·00	1·00	—
2	0·82	0·94	0·83	0·68	0·78	1·06	0·87	1·00
3	0·64	0·87	1·61	1·03	1·40	1·32	0·84	1·15
4	0·55	0·74	2·17	1·20	1·61	1·42	0·78	1·05
5	0·47	0·63	2·65	1·25	1·67	1·46	0·69	0·92
6	0·42	0·57	2·91	1·22	1·66	1·50	0·63	0·85

Table 4

Figures of merit for staggered-tuned cascades for combinations of 1 to 5 tuned circuits designed for maximally-flat amplitude characteristic. Definitions of $x_{\max g}$ and $x_{\max a}$ as for Table 3.

Number of circuits	R_t/R_0 Staggered tuned	$x_{\max g}$	$F_g = x_{\max g} \cdot R_t/R_0$ Staggered tuned	$x_{\max a}$	$F_a = x_{\max a} \cdot R_t/R_0$ Staggered tuned
1	1·00	0·35	0·35	1·00	1·00
2	0·71	0·48	0·34	1·41	1·00
3	0·50	0·78	0·39	2·00	1·00
4	0·38	1·20	0·46	2·63	1·00
5	0·31	1·46	0·45	3·23	1·00

9. Optimum Design of a Complete Amplifier

The data given in the preceding section enable a comparison to be made of the properties of designs of different numbers of circuits. For the practical design of intermediate-frequency amplifiers it is advantageous to have these data for the different possibilities that exist as regards the design of a complete amplifier. We give those data for the amplitude bandwidth between the -6 db points (amplitude fallen by a factor of 2), which is more appropriate for the complete amplifier. Cascades of 3, 4 and 5 stages are considered and for each case, all possible combinations of circuits are taken into account. For instance, with a cascade of three stages one could use 3, 4, 5 or 6 tuned circuits, because each stage can be equipped with either one or two tuned circuits. Apart from these variations further designs are possible, because, for instance, with 6 tuned circuits, all 6 tuned circuits can be combined together (according to the data of Fig. 6), or two separate combinations can be made, each with a flat group-delay curve and each with an amplitude characteristic, that is 3 db down at the limiting frequencies.

With 3 stages this would mean that one stage (with 2 tuned circuits) should have a flat group-delay curve (Fig. 2) and the other two stages (with 4 tuned circuits) together a flat group-delay curve (Fig. 4). For the single stage and for the combination of two stages, the amplitude characteristic should fall to -3 db at the limits of the frequency band considered.

From the various possibilities which thus exist, those which include a combination of only one or two tuned circuits usually have rather poor figures of merit. It is therefore necessary to combine more tuned circuits together. In Tables 6, 7 and 8, the figures of merit are listed for the most important combinations that can be made with 3, 4 and 5 stages respectively.

Table 5

Figures of merit for cascades with identical bandfilters designed for flat amplitude characteristic ($x_0^2 = 1{\cdot}4$). Definitions of $x_{\max g}$ and $x_{\max a}$ as for Table. 3.

Number of stages	R_t/R_0 Band filters	$x_{\max g}$	$F_g = x_{\max g} \cdot R_t/R_0$ Band filters	$x_{\max a}$	$F_a = x_{\max a} \cdot R_t/R_0$ Band filters
1	1·00	0·46	0·46	1·70	1·70
2	1·00	0·32	0·32	1·41	1·41
3	1·00	0·26	0·26	1·29	1·29
4	1·00	0·22	0·22	1·29	1·29
5	1·00	0·19	0·19	1·17	1·17

Table 6

Figures of merit for various possible combinations of tuned circuits in a three-stage amplifier, designed for flat group-delay characteristic. The value of $x_{\max a}$ corresponds to the -6 db points. The last line holds for a cascade of three identical bandfilters, each designed for flat amplitude.

Number of circuits	Combination	$x_{\max a} \cdot R_t/R_0$	=	F_a
3	3(III)	$1{\cdot}73 \cdot \sqrt[3]{(0{\cdot}58)^2 \cdot 0{\cdot}80}$	=	1·12
4	4(III)	$1{\cdot}96 \cdot \sqrt[3]{0{\cdot}80 \cdot (0{\cdot}67)^2}$	=	1·39
5	5(III)	$2{\cdot}04 \cdot \sqrt[3]{0{\cdot}70 \cdot 0{\cdot}66 \cdot 0{\cdot}53}$	=	1·27
5	2(I)+3(II)	$\sqrt[3]{1{\cdot}06 \cdot (1{\cdot}32)^2} \cdot \sqrt[3]{0{\cdot}94 \cdot 0{\cdot}94 \cdot 0{\cdot}80}$	=	1·10
6	6(III)	$2{\cdot}11 \cdot \sqrt[3]{0{\cdot}63 \cdot 0{\cdot}65 \cdot 0{\cdot}00}$	=	0·00
6	2(I)+4(II)	$\sqrt[3]{1{\cdot}06 \cdot (1{\cdot}41)^2} \cdot \sqrt[3]{0{\cdot}94 \cdot 0{\cdot}80 \cdot 0{\cdot}69}$	=	1·03
6	Three identical flat-amplitude band filters ($x_0^2 = 1{\cdot}4$)			1·49

In these tables, the combinations are indicated by an arabic figure followed by a roman figure between brackets. The arabic figure gives the number of tuned circuits that are combined together to give a flat delay curve; the roman figure gives the number of stages which contain this combination. Thus 4(II) +5(III) means that there are 5 stages in total (II+III), of which two stages with 4 tuned circuits are combined together, and the other three stages which contain 5 tuned circuits are also combined together.

Due to the use of combinations of more than two tuned circuits, where the amplitude bandwidth is always the limiting factor (see Table 3), we have to use the F_a—figure of merit. In Tables 6-8, this F_a value is first written as a product of two factors. The first factor gives the mean value of $x_{\max a}$, the second factor gives the mean value of R_t/R_0. The values of $x_{\max a}$ are derived from the curves of Figs 1-6 by reading the value of x at the -6 db or -3 db points. The data for R_t/R_0 are taken from the insets in Figs. 1-6; where different possibilities existed for one combination, the data corresponding to the largest gain were chosen.

The data given in the last line of the tables enable a comparison with the figure of merit of an amplifier containing band filters and designed for flat amplitude characteristic only. For the ratio of this figure of merit to the best figure of merit for the flat group-delay designs one thus finds

for three stages $1{\cdot}49/1{\cdot}39 = 1{\cdot}07$,

for four stages $1{\cdot}41/1{\cdot}19 = 1{\cdot}18$,

for five stages $1{\cdot}35/1{\cdot}05 = 1{\cdot}29$.

The price to be paid for an amplification with neither amplitude nor phase distortion is thus a total loss in gain of $(1{\cdot}07)^3 = 1.2$; $(1{\cdot}18)^4 = 1{\cdot}9$ and $(1{\cdot}29)^5 = 3{\cdot}5$ for the three cascades considered. When judging this consequence of a distortion-free amplifier one should take into account the fact that a detector stage should not be counted as an i.f. stage.[7]

An amplifier containing a convertor valve and three amplifier valves followed by a diode detector thus comprises three i.f. stages proper.

From Tables 6-8 it follows that for each number of stages, the designer has the possibility of choosing between various numbers of circuits; the differences in gain are in many cases only small. This makes it possible to take other practical considerations into account, for instance the wish to use a large number of circuits so as to increase the selectivity, or the possibility of dividing the i.f. amplifier into two identical sections, each a combination with flat delay characteristic. The adjustment procedure is in this case somewhat easier than if all circuits are combined together.

Table 7

Analogous data as in Table 6, but for a four-stage amplifier.

Number of circuits	Combination	$x_{\max a} \cdot R_t/R_0$	=	F_a
4	4(IV)	$1{\cdot}96 \, . \sqrt[4]{(0{\cdot}45)^2 \, . \, (0{\cdot}67)^2}$	=	1·08
5	5(IV)	$2{\cdot}04 \, . \sqrt[4]{0{\cdot}70 \, . \, (0{\cdot}56)^2 \, . \, 0{\cdot}53}$	=	1·19
6	6(IV)	$2.11 \, . \sqrt[4]{0{\cdot}63 \, . \, 0{\cdot}65 \, . \, (0{\cdot}50)^2}$	=	1·19
6	2(II)+4(II)	$\sqrt{1{\cdot}06 \, . \, 1{\cdot}41} \, . \sqrt[4]{(0{\cdot}82)^2 \, . \, 0{\cdot}80 \, . \, 0{\cdot}69}$	=	0·95
6	3(II)+3(II)	$\sqrt{(1{\cdot}32)^2} \, . \sqrt[4]{(0{\cdot}94)^2 \, . \, (0{\cdot}80)^2}$	=	1·15
7	3(II)+4(II)	$\sqrt{1{\cdot}32 \, . \, 1{\cdot}41} \, . \sqrt[4]{0{\cdot}94 \, . \, 0{\cdot}80 \, . \, 0{\cdot}80 \, . \, 0{\cdot}69}$	=	1·10
8	4(II)+4(II)	$\sqrt{(1{\cdot}41)^2} \, . \sqrt[4]{(0{\cdot}80)^2 \, . \, (0{\cdot}69)^2}$	=	1·04
8	Four identical flat-amplitude band filters ($x_0^2 = 1{\cdot}4$)			1·41

The average gain per stage follows directly from the figure of merit F_a, because

$$F_a = x_{\max a} \cdot \frac{R_t}{R_0} = 2\Delta\omega_{\max a} R_0 C \cdot \frac{R_t}{R_0} = $$
$$= 2\Delta\omega_{\max a} R_t C.$$

With a mutual conductance g_m, the gain per stage is given by

$$g_m R_t = \frac{F_a}{2\pi\Delta f C}$$

where Δf is the width of the frequency band between the -6 db points and C is the total capacitance per stage.

We illustrate this result with a practical case of a four-stage i.f. amplifier.

We chose in Table 7 the 3(II)+3(II) combination. With $g_m = 6{\cdot}5$ mA/V and $C = 25$ pF, the average voltage gain per stage will be

19 for a bandwidth of 2·5 Mc/s

11 for a bandwidth of 4·5 Mc/s.

The i.f. amplifier contains two stages with identical band filters and two stages with a single tuned circuit. The parallel resistance of the tuned circuits of the band filters (presuming equal primary and secondary capacitances) follows from

$$R = \frac{x_{\max a}}{2\Delta\omega_{\max a} \cdot \frac{1}{2}C} = \frac{1{\cdot}32}{\pi\Delta f C}$$

(Δf = bandwidth between the -6 db points), and the coupling coefficient is given by

$$kQ = x_0 = 1{\cdot}41.$$

With unequal primary and secondary capacitances, the parallel resistances should in the usual way be adapted to the capacitances.

The parallel resistance of each of the single tuned circuits follows from

$$R_1 = \frac{\tau_1}{\tau_0} \, . \, R_0 = \frac{\tau_1}{\tau_0} \cdot \frac{x_{\max a}}{2\Delta\omega_{\max a} C} = \frac{0{\cdot}80 \, . \, 1{\cdot}32}{2\pi\Delta f C}$$

with Δf = bandwidth between the -6 db points.

10. The Adjustment of Phase-linear I.F. Amplifiers

The tuning and damping of all tuned circuits and the coupling coefficients of the band filters can be derived from the data given in the preceding sections. The adjustment can follow the same procedure as with flat-amplitude-characteristic amplifiers: the damping resistances and the coupling coefficients of the band filters are given the correct values; subsequently each circuit is tuned in the conventional way to its prescribed resonance frequency using an amplitude measurement as indication.

The accuracy and reproducibility of such an adjustment is exactly the same as that of a corresponding adjustment of a flat-amplitude-characteristic amplifier. If the accuracy of this method is considered sufficient for amplifiers with flat amplitude curves, there is no need whatsoever to introduce group-delay measurements in the adjustment procedure of phase-linear amplifiers.

Circumstances are somewhat different during the original design and development period.

Table 8

Analogous data as in Table 6, but for a five-stage amplifier.

Number of circuits	Combination	$x_{\max a} \,.\, Rt/R_0$	=	F_a
5	5(V)	$2{\cdot}04 \,.\, \sqrt[5]{(0{\cdot}38)^2 \,.\, (0{\cdot}56)^2 \,.\, 0{\cdot}53}$	=	0·96
6	6(V)	$2{\cdot}11 \,.\, \sqrt[5]{0{\cdot}63 \,.\, (0{\cdot}44)^2 \,.\, (0{\cdot}50)^2}$	=	1·05
6	3(II)+3(III)	$\sqrt[5]{(1{\cdot}32)^2 \,.\, (1{\cdot}32)^3} \,.\, \sqrt[5]{0{\cdot}94 \,.\, 0{\cdot}80 \,.\, (0{\cdot}58)^2 \,.\, (0{\cdot}80)}$	=	0·96
7	3(III)+4(II)	$\sqrt[5]{(1{\cdot}32)^3 \,.\, (1{\cdot}42)^2} \cdot \sqrt[5]{(0{\cdot}58)^2 \,.\, 0{\cdot}80 \,.\, 0{\cdot}80 \,.\, 0{\cdot}69}$	=	0·93
7	4(III)+3(II)	$\sqrt[5]{(1{\cdot}32)^3 \,.\, (1{\cdot}42)^2} \,.\, \sqrt[5]{0{\cdot}80 \,.\, (0{\cdot}67)^2 \,.\, 0{\cdot}94 \,.\, 0{\cdot}80}$	=	1·05
8	3(II)+5(III)	$\sqrt[5]{(1{\cdot}32)^2 \,.\, (1{\cdot}46)^3} \,.\, \sqrt[5]{0{\cdot}94 \,.\, 0{\cdot}80 \,.\, 0{\cdot}70 \,.\, 0{\cdot}66 \,.\, 0{\cdot}53}$	=	1·00
8	4(II)+4(III)	$\sqrt[5]{(1{\cdot}42)^2 \,.\, (1{\cdot}42)^3} \,.\, \sqrt[5]{0{\cdot}80 \,.\, 0{\cdot}69 \,.\, 0{\cdot}80 \,.\, (0{\cdot}67)^2}$	=	1·02
9	5(III)+4(II)	$\sqrt{(1{\cdot}46)^3 \,.\, (1{\cdot}42)^2} \,.\, \sqrt[5]{0{\cdot}70 \,.\, 0{\cdot}66 \,.\, 0{\cdot}53 \,.\, 0{\cdot}80 \,.\, 0{\cdot}69}$	=	0·96
10	4(II)+6(III)	$\sqrt{(1{\cdot}42)^2 \,.\, (1{\cdot}50^3} \,.\, \sqrt[5]{0{\cdot}80 \,.\, 0{\cdot}69 \,.\, 0{\cdot}63 \,.\, 0{\cdot}65 \,.\, 0{\cdot}00}$	=	0·00
10	Five identical flat-amplitude band filters ($x_0^2 = 1{\cdot}4$)			1·35

It is certainly possible to develop a phase-linear amplifier, checking the result by accurately measuring the amplitude curve and comparing this curve with the prescribed characteristic. Possible group-delay deviations will be small as compared with the group-delay variations that are always present in the conventional amplifiers with flat amplitude characteristics.

However, a direct measurement of the group-delay characteristic is obviously a great help. In our experimental work on phase-linear television receivers[5] we used a "wobbulator" measurement equipment[7], which allows the direct tracing of both group-delay and amplitude characteristics. It soon became our practice to consider mainly the group-delay curve, which measurement, as it is much more sensitive than an amplitude measurement, enables a more accurate adjustment to be made. Amplitude measurements were made occasionally to check the correct place of the − 6 db point in the frequency spectrum.

According to our experience it is quite feasible to adjust the tuning of the circuits of either a phase-linear or a flat-ampltitude i.f. amplifier, using group-delay measurements as references. The accuracy of such a tuning procedure is better than can be achieved by amplitude measurements, and therefore the reproducibility of the characteristics of amplifiers which are trimmed in this way will also be better.

11. References

1. J. Harmans, "On the design of wide-band amplifiers," *Funk und Ton*, No. 2, pp. 72-80, February 1948.
2. A. Fromageot and P. Belgadère, "Determination of the elements of an amplifier with circuits tuned to different frequencies and such that the phase-frequency function is linear in a given range of frequencies," *Annales de Télécommunication*, **4**, pp. 169-176, May 1949.
3. J. Laplume, "Intermediate frequency amplifier of reduced phase distortion," *L'Onde Electrique*, **31**, pp. 357-362, August-September 1951.
4. J. H. de Boer and A. van Weel, "An instrument for measuring group delay," *Philips Technical Review*, **15**, pp. 307-316, May 1953.
5. A. van Weel, "Phase linearity of television receivers," *Philips Technical Review*, **18**, pp. 33-51, July 1956.
6. A. van Weel, "Some remarks on the radio frequency phase and amplitude characteristics of television receivers," *J.Brit.I.R.E.*, **16**, pp. 271-280, May 1956.
7. A. van Weel, "Analysis and design of diode detector stages," to be published.
8. C. J. Heuvelman and A. van Weel, "Group delay measurements," *Wireless Engineer*, **33**, pp. 107-113, May 1956.

RECENT DEVELOPMENTS IN MOBILE RADIO IN BRITAIN*

by

John R. Brinkley (Member)†

SUMMARY

The paper outlines briefly the development of v.h.f. mobile radio in Great Britain, its past history, its present development, and its future planning. The predominant use of amplitude modulation is discussed, particularly in relation to the introduction of 25 kc/s channel spacing which is now being used. Equipment has also been developed to work on channel spacings as close as 12.5 or 15 kc/s. Future trends and the possibility of introducing single-sideband with or without carrier, are briefly discussed.

1. Introduction—Historical Background

The development of mobile radio in Britain has followed a pattern which has marked similarities and marked differences from the development in the United States. In order to appreciate recent trends it is necessary first to trace briefly the history of British mobile radio development.

Before the war, mobile radio in Britain was confined almost exclusively to the Police using medium frequency (2 Mc/s) transmissions. A number of police forces established experimental v.h.f. services about 1935 in the 80 and 95 Mc/s bands, but progress was impeded by long delays in allocating permanent bands to mobile radio. Eventually at the 1938 Cairo Convention, three bands were allocated, 80–84, 95–100, and 128–132 Mc/s. These bands were subsequently modified.

During the war, land mobile development was confined to Police, Fire and Civil Defence, and it is interesting to note that v.h.f. apparatus designed by a British Company in 1938-9 for the Police radio bands became the basis of all Allied ground-to-air v.h.f., which was a war-winning development complementary in importance to radar. As a civil defence measure a countrywide survey was carried out in every major city and county in the U.K. Based on this survey an extensive network of stations was built for police, fire and civil defence, until today about 300 fixed stations and 3,000 mobiles are in operation for these services. All but a very few remote areas have excellent police and fire service coverage.

At the end of the war demand arose for the development of general mobile services and in 1947 the first taxi scheme was established in Cambridge. In the same year the first v.h.f. marine service was established for a tug company on the River Tyne.

Since that date there has been a steady development of mobile radio until today there are over 2,000 fixed and about 20,000 mobile stations operating on the v.h.f. bands in the following categories:

(*a*) Police—Fire
(*b*) Taxi
(*c*) Marine
(*d*) Electricity, Fuel and Power
(*e*) General Commercial Services
(*f*) Industrial Services
(*g*) Airfield Mobiles

Of the 20,000 mobiles, more than 19,000 employ amplitude modulation, and less than 1,000 use frequency modulation. This is a radical difference from U.S. practice, which will be discussed later.

2. Present Development

The present stage of development wi.l be described under various headings.

* Manuscript received 22nd March 1957. This paper was presented in Session 2 of the 1957 National Convention of the Institute of Radio Engineers in New York on the 18th March, and is published here by permission of the Institute. The paper will subsequently appear in Part 8 of the 1957 *I.R.E. National Convention Record*. (Paper No. 398.)

† Pye Telecommunications Ltd., Cambridge.

U.D.C. No. 621.396.931.

2.1. *Administration*

The administration and frequency allocation for mobile radio are the responsibility of the British Post Office.

This Department operates the telephone and other competitive communications services, and there has been a good deal of pressure to set up an independent allocation authority similar to the F.C.C. which many people feel would be preferable. A Mobile Radio Advisory Committee has been set up which takes the view of users and their associations, manufacturers and expert advisers, and makes policy recommendations to the Postmaster-General. This committee has made two reports. The first[1] is concerned with the loss of land mobile frequencies to Commercial Television, and the second[2] deals with the introduction of narrow (split) channelling.

2.2. *Frequency Allocation*

The frequency bands now available to land and marine mobile radio are as follows:

General Land Mobile	71·5 – 72·8 Mc/s
	76·95– 78·0 Mc/s
	85·0 – 86·7 Mc/s
	86·95– 88·0 Mc/s
Police and Fire ...	80·0 – 84·0 Mc/s
	95·0 –100 Mc/s
Marine	156–165 Mc/s
Land Mobile	165–173 Mc/s
Land Mobile (tentatively allocated and under review) ...	460–470 Mc/s

The bands between 70 and 100 Mc/s have proved very suitable and are probably optimum in frequency for mobile services. There is no 25–50 Mc/s band in the U.K. and skip problems and interference from other countries are virtually unknown.

It will be seen that there is a large marine band in the U.K. This large allocation scarcely reflects need and many of the channels are unused.

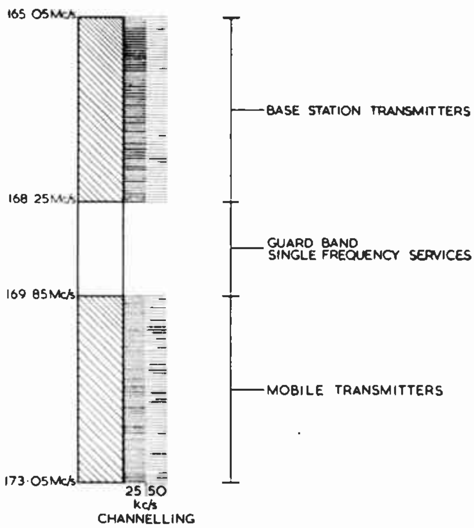
2.3. *Double Frequency Working*

Nearly all allocations are on a two-frequency basis (separate "go" and "return" frequencies with go and return blocks widely separated). This has led to a minimum of interscheme interference and it has been possible to avoid almost entirely the severe problems of intermodulation so common in the U.S. which are aggravated by single frequency allocation. In this respect at least, mobile radio frequency allocation appears to be on a far more satisfactory basis in Great Britain.

Table 1 shows the layout of frequencies in the band 165–173 Mc/s. It will be seen that all fixed stations transmit on 165·05–168·25 Mc/s and all mobiles on 169·85–173·05 Mc/s, the spacing between go and return frequencies being constant. A guard band of 1·6 Mc/s is left between the go and return blocks since if this were not done, single frequency allocation conditions would exist on those channels where go and return spacings was closer than about 1·6 Mc/s. Channels in the guard band are also allocated, but with geographic restrictions imposed on allocation.

Table 1

United Kingdom Channelling Plan
165–173 Mc/s



With the arrangement shown, it is possible to obtain 64 double frequency channels (50 kc/s) or 128 double frequency channels (25 kc/s) which are all largely free from mutual interference. This channel utilization could never be achieved with single frequency working, where in spite of the apparent doubling of channels, far fewer interference-free channels could be obtained in any one area.

The double frequency method of allocation gives greater freedom in siting, and in this respect it greatly eases the work of both the scheme planner and the frequency allocator.

Double frequency allocation also has the great advantage that frequencies can be repeated geographically at much closer spacings than is possible with single frequency allocation.

The important conclusion to be drawn from British experience of frequency allocation is that provided double frequency allocation is used, closer and closer channel spacings will make more and more channels available without multiplying interference problems. If single frequency working were adhered to, closer channel spacing would have much more limited benefits because of the large number of interference products which would accrue.

2.4. *Channelling Standards*

Until recently channels were spaced evenly 100 kc/s apart in the 156–173 Mc/s band, and 50 kc/s apart in the 70 Mc/s band. This has now been reduced to 50 kc/s in the 156–173 Mc/s band.

The introduction of 25 kc/s channelling in the 170 Mc/s and 70 Mc/s bands is under consideration and in anticipation of its adoption most apparatus now being installed will effectively separate channels spaced 25 kc/s apart on both bands. This is leading to a position where some 524 channels are potentially available for general development and 360 channels for Marine. As a result, the shortage of frequencies has been greatly eased in the congested area (Greater London) and scarcely exists any longer in the rest of the country. This is a situation which is no doubt viewed with some envy in the U.S.

This easement of channel crowding will lead to a slower development of the 450 Mc/s band in Great Britain than is likely in the U.S. and Canada.

Recent tests have shown that channels spaced as close as 12·5 kc/s may be effectively separated without undue complication, and it is likely that some trial schemes will be installed using this spacing, particularly in the 70 Mc/s band. The majority of the technical progress in the very narrow channels has been made to date using amplitude modulation.

3. Modulation System

The most outstanding difference between U.S. and British mobile radio practice has been the dominant use of amplitude modulation in Britain, and the dominant use of frequency modulation in the U.S. Thus in Britain, of over 22,000 equipments, less than 1,000 are f.m. and of these less than 100 are in the commercial category. The reasons for this wide diversion of practice between the two countries are rewarding to study, if this is done impartially.

At the time when the U.S. was changing to f.m. the British had already a substantial number of a.m. equipments in use. It was found that by the addition of good noise limiters to the a.m. equipment, suppression of ignition noise, by far the most important noise source, was about equally effective in the two systems. As a result of this and other factors, e.g., general simplicity and greater frequency tolerance, a preference for a.m. grew up. It was further felt even at that stage, that as channel spacing grew closer a.m. might prove preferable. In the immediate post-war period, however, the British Companies were manufacturing both a.m. and f.m. equipment, but over the years a.m. has become dominant, and some f.m. installations have actually been recently replaced by a.m.

The author would not attempt to draw emphatic conclusions from this state of affairs beyond stating that there can be no doubt that a.m. mobile systems work very well. The two systems are different and one is not necessarily better than the other. As channel spacings close in, however, comparisons between the two systems take on a new importance.

4. Modulation System and Channel Spacing

The early controversy between a.m. and f.m. for mobile systems was largely concerned with noise and centred mainly around the superior suppression of thermal noise achieved with wide deviation f.m. systems. Unfortunately at the time of the introduction of f.m. on a wide scale in the U.S., the tremendous development of mobile radio was not foreseen, and it is possible that insufficient consideration was given to the position that would arise when very narrow channelling had to be introduced. The problem in U.S. mobile radio at this present moment is not thermal noise suppression but verbal suppression. An extreme example of this is the Boston No. 1 Taxi Channel, on which no fewer than 17 taxi companies may attempt to communicate simultaneously. The modulation system which is important under such conditions is the one which may readily give the greatest number of separate channels.

There are two sets of comparisons which now have to be made. The first is between a.m. and f.m. in their present stage of development, and the second is the longer term comparisons of all the systems which can be put forward with a view to achieving the ultimate in close channel spacings.

Comparing a.m. and f.m. in the first case, it must be remembered that there is nothing static about the development of both systems. Current British experience shows that it is relatively easy to design and manufacture a.m. equipment capable of separating very adequately spacings as close as 12·5 or 15 kc/s. It is probable that with only a little more development 10 kc/s channelling will be shown to be practicable. Work on similar channel spacings for f.m. is not so advanced and is more difficult, and firm conclusions about the limitations on f.m. channel spacings have not

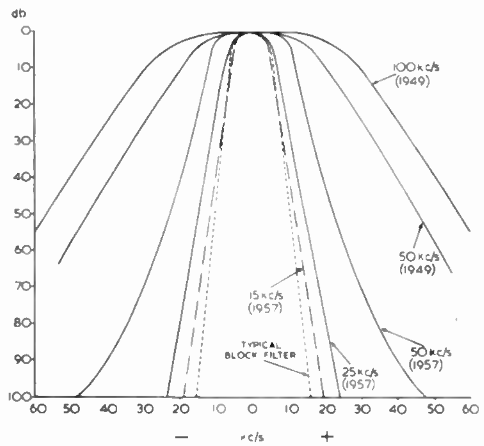


Fig. 1. Progressive increase in selectivity of i.f. amplifiers, 1948-57.

yet been drawn. There can be little doubt, however, that for simplicity of filter design and tolerance to residual drift, a.m. has significant advantages. Fig. 1 shows the i.f. amplifier responses used in typical a.m. receivers, including the response required for channel spacings as close at 12·5 kc/s.

It has been suggested that with a.m. spurious sideband radiation would prove as troublesome or more troublesome than f.m. when very close channel spacings were employed. This has not been found to be the case, provided simple precautions including limiting and filtering are taken.

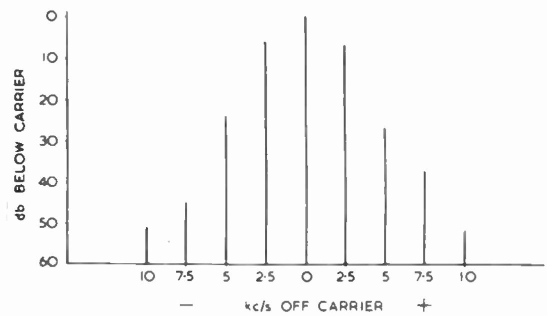


Fig. 2. Spurious sideband radiation from 15 watt a.m. transmitter (with clipper and low pass filter fitted). Carrier frequency: 162·7 Mc/s. Modulation frequency: 2·5 kc/s. Modulation level: +16 db above 50% modulation.

Figure 2 shows the spurious sideband radiation from a typical 15 watt a.m. transmitter. The transmitter under test, a production unit, shows components at 10 kc/s off tune more than 50 db down with modulation +16 db above 50 per cent. modulation. This result is obtained with very simple techniques of limiting and filtering. Improved techniques including perhaps envelope feedback would achieve even better results.

The longer term assessment of close channelling requires the consideration of four systems: a.m., f.m., single sideband a.m., and single sideband with carrier transmitted (compatible s.s.b.). It is possible that further systems or variations may yet be proposed.

Table 2 shows the existing and probable capabilities of the four systems so far as they are known with C.C.I.R. line telephony s.s.b. taken as a datum (0·3–3·4 kc/s, spaced at 4 kc/s intervals).

It will be seen that the ultimate spacing for a.m. d.s.b. is likely to be in the region of 8–10 kc/s which is comparable with the present medium-wave broadcast band The ultimate spacing for f.m. is yet to be established but may well be considerably greater.

Mobile s.s.b. with or without carrier will be limited to 4–5 kc/s and it may be very difficult

Table 2

Channel Spacings with Different Modulation Systems

System	Channel Spacing: Currently in Use	Now Possible	Eventual Development
Line S.S.B.	4 kc/s	4 kc/s	4 kc/s
A.M. D.S.B.	25 kc/s	12·5 kc/s	8–10 kc/s
F.M.	30 kc/s	20 kc/s	?
Mobile S.S.B. with carrier	—	—	4·5 kc/s

in practice to achieve these figures. S.s.b. for mobile radio would appear therefore to have a theoretical 2:1 advantage over d.s.b. a.m. but whether this advantage could be achieved and maintained in practice is not so certain, and it is as yet uncertain where f.m. stands in the ultimate picture.

The complexity of s.s.b. for v.h.f. mobile systems is, of course, very great. The system of compatible s.s.b. described by Kahn,[3,4] in

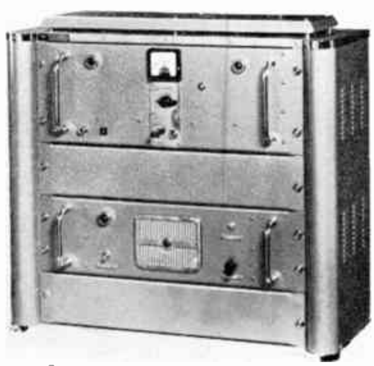


Fig. 3. Fixed station equipment.

which full carrier is radiated with one sideband, has obvious attractions not only for its simplicity of reception, but because of the degree of compatibility with d.s.b. a.m. One of the long-term disadvantages of f.m. is that if change is eventually to be made to s.s.b. in order to achieve the ultimate channel spacing, each f.m. system will become obsolete at one stroke, whereas it is possible to envisage a stage-by-stage changeover from d.s.b. to s.s.b. with carrier. In view of the vast amounts of equipment likely to be operating when such changeovers are contemplated, this point could eventually be one of considerable importance. This stage has, in fact, now been reached in the air-to-ground h.f. telephony field where a changeover from d.s.b. a.m. to either s.s.b. or s.s.b. with carrier is being considered.

5. Marine Development

Much consideration has recently been given in Britain and other countries to the standardization of v.h.f. for international marine services. British shipping used a.m. v.h.f. and would have continued to do so but for the fact that the U.S. was committed to f.m. by the Atlantic City Agreement of 1947. Since the conversion to f.m. was agreed in order to ensure common standards with the U.S. shipping, it was also very wisely agreed to adopt U.S. standards more or less *in toto*.

The frequency modulation specification finally agreed by the British Post Office[5], and British industry, is closely in line with F.C.C., U.S., R.E.T.M.A., Canadian D.O.T. and the recent Hague Conference recommendations. The standards should permit the introduction of 50 kc/s channelling for Marine Services, and it is to be hoped that after years of conferences and controversy international marine v.h.f. will now develop in a healthy manner.

The value of international standardization in the field of communication in which techniques are developing rapidly is open to question and the marine experiment to this end must be watched. An extension of international standardization thinking to the land mobile fields, where it would appear largely superfluous, or even embarrassing would, in the author's view, be an unwelcome impediment to progress at this stage, since it could prevent each country introducing new techniques at the rate most suited to its technical and economic development.

6. Equipment

Due to licensing conditions and shorter range requirements (a maximum of 25 watts for commercial frequencies) British v.h.f. equipment has tended to be lower in power than the U.S. equivalent. Transmitters range from 3 to 20 watts, and the use of very compact under-the-dash mobiles of about 3–6 watts output is

Fig. 4. Dash mounting mobile equipment.

very common. By using the receiver l.f. output as modulator the number of valves can be kept down to as low as 12. Current consumptions are low and heavy-duty dynamos are seldom necessary.

The new land mobile specifications have necessitated the addition of audio-frequency limiters and low-pass filters in the transmitters. The audio filters are inserted at the modulator output in order to minimize spurious sideband radiation. The filters which in the a.m. case operate at high level must be air-cored, since a magnetic core material is found to introduce more sideband harmonics than the filter removes. If these simple precautions are taken, unwanted sideband radiation is largely undetectable even at 12·5 kc/s spacing.

The a.m. receivers are now very similar to their U.S. f.m. equivalent. I.f. filter design and frequency stability are not so stringent in the a.m. case.

Figures 3, 4, 5 show typical fixed and mobile stations with outputs as follows:

Fixed: 15 W a.m./20 W f.m.

Dash mounting mobile: 5 W a.m./10 W f.m.

Boot mounting mobile: 15 W a.m./20 W f.m.

Different variations of the equipment are

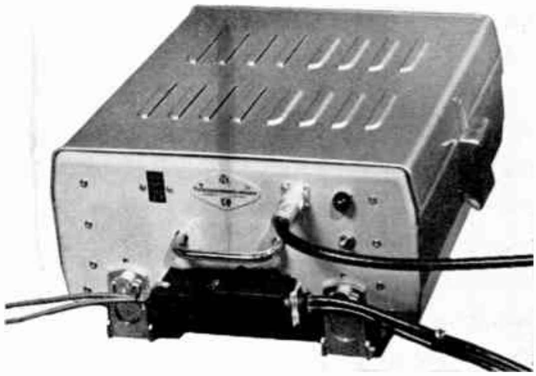


Fig. 5. Boot mounting mobile equipment.

offered for 100 kc/s (aeronautical), 50–60 kc/s and 20–25 kc/s working. 12·5–15 kc/s channelling has not yet been offered commercially. The f.m. versions are offered mainly for marine and overseas markets.

In the development of 12·5 kc/s equipment consideration is being given to the use of quartz crystal i.f. filters which appear to be most effective. A block diagram of such a receiver is shown in Fig. 6 and a practical receiver using a quartz crystal filter of U.S. design is shown in Fig. 7.

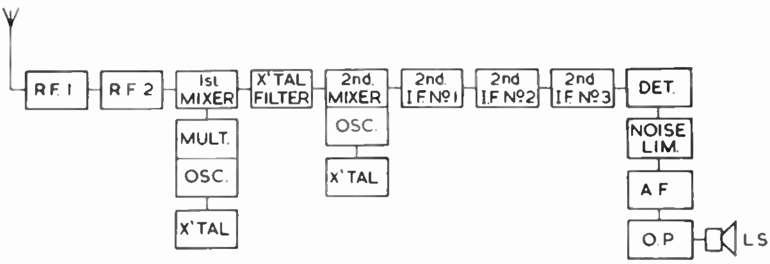


Fig. 6. Block schematic of a.m. receiver for 12·5–15 kc/s channelling using crystal i.f. filter.

7. Conclusions

The development of mobile equipment in Britain has been on somewhat different lines to the U.S. pattern. The main difference has been the widespread use of amplitude modulation as opposed to frequency modulation in the U.S.

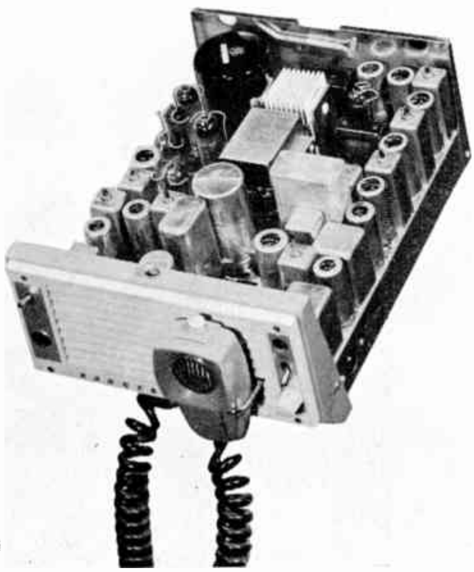


Fig. 7. Dash mounting mobile equipment incorporating a quartz crystal filter for 12·5 kc/s channelling.

Experience in Britain has shown that a.m. gives an excellent performance for mobile services. It is readily suitable for narrow channel working and may well prove to have important advantages in this respect in the long term.

8. References

1. "Report of the Mobile Radio Committee." (Her Majesty's Stationery Office, London, 1955.)
2. "Second Report of the Mobile Radio Committee." (Her Majesty's Stationery Office, London, 1956.)
3. Leonard R. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proceedings of the Institute of Radio Engineers*, **40**, No. 7, July 1952.
4. Leonard R. Kahn, "A Compatible Single Sideband System," I.R.E. Second Annual Symposium on Aeronautical Communications, 1956.
5. Post Office Engineering Department Specification No. W.6188A, "Private Metric (V.H.F.) Mobile Radio Telephone Services. Specification for Angle Modulated Transmitter-Receivers (Fixed and Mobile). For 50 kc/s carrier Frequency Separation and Maximum Deviation up to ± 9 kc/s."

# of current interest . . .

Abbreviations Translated

Reference to the confusion often caused by the wide use of abbreviations was made in the Editorial to the April *Journal*. Firms and organizations in the telecommunications and electronics industries, both in Great Britain and elsewhere, seem prone to using abbreviations to describe themselves, and a glossary of these abbreviations which will prove a boon to the uninitiated has been prepared by the technical librarian of Marconi's Baddow Research Laboratories, Mr. S. T. Cope. Entitled "Glossary of Abbreviations for names of technical, scientific, industrial and professional organizations, with particular reference to the telecommunications industry," it is the second edition of a booklet first published two years ago. Copies may be obtained from Marconi's W.T. Co. Ltd., Chelmsford, price 2s. 6d.

I.T.A. Service for Southern England

The Independent Television Authority is to open its seventh transmitting station in the late Spring of 1958. This will provide a service for an area containing more than two million people in Southern England. The service area is roughly a half circle, stretching along its base from Weymouth through Ventnor to Brighton and reaching Newbury in the north. It will thus cover Hampshire and the Isle of Wight, almost the whole of Dorset, West Sussex, and south-eastern parts of Wiltshire and the south-western parts of Surrey.

With the opening of this station in the Isle of Wight the Authority's programmes will be available to 80 per cent. of the total population of the United Kingdom. Coverage of a further 5 per cent. of the population will be added in 1958 with the opening of a station to serve the north-east coast.

B.B.C. Television and V.H.F. Stations

The B.B.C.'s television station for West Wales at Blaen Plwy, near Aberystwyth, came into service on 29th April last. On the same day the v.h.f. f.m. service, previously operating with one programme on low power with a temporary mast and aerial, was brought into full power operation in conjunction with the new 500 ft. mast which supports both the television and the v.h.f. broadcasting aerials.

The television transmissions will be on Channel 3 (vision 56·75 Mc/s, sound 53·25 Mc/s), horizontally polarized, and with an effective radiated power of 1 kW. They will serve the coastal belt around Cardigan Bay including the Pwllheli and Portmadoc areas in the north and the Tregaron and Aberporth areas in the south. The v.h.f. sound broadcasts are on Band II on 93·1 Mc/s, 88·7 Mc/s and 90·0 Mc/s for the Welsh Home Service, Light Programme and Third Programme respectively. Each programme is transmitted with an effective radiated power of 60 kW.

The v.h.f. radio stations associated with the television transmitter at Norwich and Sutton Coldfield, previously operating on reduced power, have been operating on full power since April 30th. Each station has three transmitters for the three programmes, each having an effective radiated power of 120 kW.

Mobile Radio—25 kc/s Channelling

The British Post Office is making arrangements for trials of mobile radio equipment using 25 kc/s channelling instead of the 50 kc/s channelling equipment so far authorized.

In the meantime frequency allocations for private mobile radio services are to be made only for 50 kc/s channel spacing in both high and low bands and equipment used must, therefore, meet the minimum technical requirements of the 50 kc/s specifications which became effective from January 1st, 1957.

These statements are made in a letter to the Radio Communication and Electronic Engineering Association which had asked the G.P.O. for clarification of the present position.

The Post Office adds that while no guarantee can be given, as there are so many unforeseeable factors, it is the general policy to arrange matters so that users' equipments may run an economic life. "When the next stage is reached in the introduction of narrower channelling standards for mobile services," the letter states, "the Mobile Radio Committee will no doubt make a specific recommendation about the change-over period, as they did for the change from 100 kc/s to 50 kc/s in the high band."

. . . Radio Engineering Overseas

A selection of abstracts from European and Commonwealth journals received in the Library of the Institution. All papers are in the language of the country of origin of the journal unless otherwise stated. The Institution regrets that translations cannot be supplied.

621.314.7

Limit ratings in transistor operation. M. GASCHI. *Onde Electrique*, **37**, pp. 239-243, March 1957.

The paper examines the basic factors limiting transistor applications. After a survey on the factors limiting the maximum applicable current and voltage. examination is made of the conditions which limit the maximum applicable power. Due to their sensitivity to temperature effects. transistors could meet overheating conditions leading to run-away. After analysing the condition for stable operation. a method is described from which curves can be drawn. permitting a choice of the operating point, together with the knowledge of the type of operation and forms of applied signals.

621.314.7

Power transistors. C. DUGAS and J. GROSVALET. *Onde Electrique*, **37**, pp. 231-234. March 1957.

The maximum power which can be dissipated by a transistor is considered. From this a working area is deduced which is dependent on parameters determined by the conditions of use of the transistor. The influence of the physical characteristics on the electrical properties of the semi-conductor is then considered. and in particular the variation of these parameters, especially gain. when operating at high current values. From these considerations general principles emerge which must be satisfied by manufacturing techniques.

621.316.8:621.317.34

Measurement of noise in resistors. J. GAJSEK. *Elektrotehniski Vestnik*, **25**, pp. 29-32. January/February 1957.

Noise in resistors is due not only to temperature but also to current effects. The latter are important in carbon layer resistors. Unlike thermal noise this noise cannot be calculated, but must be measured. An apparatus is described which can also measure background noise on high quality resistances. i.e. to 0·25 μV/V. Resistors of 5 kΩ and more can be measured by this device if loaded up to 500 V.d.c. The device is composed of the following main parts: input circuit with standard resistors from 5 kΩ to 1 MΩ. amplifier for obtaining 110 db gain with the least possible noise and hum, indicator with loudspeaker, electronically stabilized rectifier for amplifier supply, ganging oscillator with the attenuator and the resistor-loading device.

621.317.42

Special application of instruments for measuring intensity of field. L. HINTZBERGEN. *Elektrotehniski Vestnik*, **25**, pp. 7-12. January/February 1957.

For the determination of power and the best position of a transmitter, whether from the geographical or electrical point of view, it is necessary to know the conductivity of the soil near the transmitter. The paper describes instruments for measuring intensity of field with the object of determining the conductivity. Propagation theory is discussed and several practical examples cited.

621.317.733

Direct indicating bridge for the measurement of temperature rises in electrical appliances. F. KAMMER. *Proc. Instn Radio Engrs, Aust.*, **18**, pp. 81-87, March 1975.

An instrument is described for use in the routine measurement of the temperature rise, under load, in the windings of electrical appliances. It consists of a direct reading Wheatstone bridge compensated for changes in ambient temperature. A special connector box is provided to facilitate rapid measurement.

621.319.4

Quality control in radio component manufacturing. M. LEPRINCE-RINGUET. *Onde Electrique*, **37**, pp. 301-302. March 1957.

The paper describes the various quality controls applied in the manufacture of mica capacitors. Quality control begins with the selection of material, is operative through all stages of production and covers individual component testing after manufacturing. Component testing is the responsibility of the production department. To provide even tighter control, a final "super-control" applies batch-testing methods to component samples before they leave the factory. Internal specifications define the various test procedures and determine the sampling methods (statistical or percentage) and tolerance figures.

621.372.8

Reflection at the junction of a rectangular waveguide and sector horn. G. PIEFKE. *Archiv der Elektrischen Ubertragung*, **11**, pp. 123-135, March 1957.

The reflection coefficient is calculated at the junction of a rectangular waveguide and a "sector horn." The horn is flared once in the H-plane (H-case), once in the E-plane (E-case) of the incoming TE_{10} mode. The reflected modes are in the H-case the TE_{n0} with $n=1, 3, 5$, etc., in the E-case the modes TE_{1n} and TM_{1n} with $n=0, 2, 4$, etc. The range $1 \leqq 2a/\lambda_0 \leqq 2$ (a width of the waveguide, λ_0 wavelength in free space) the reflection coefficient is represented in terms of magnitude and phase by curves for flare angles up to 90°. In the H-case, but not in the E-case, the reflection coefficient is independent of the height b of the waveguide. For this reason $a=2b$ has been assumed with the curves of the E-case. For very small flare angles an approximation formula is also given. The terminating impedance presented by the horn of infinite length at the junction between waveguide and horn is a resistance shunted in the H-case by an inductance, in the E-case by a capacitance. Most of the resistance is due to the wave impedance of the TE_{10} mode. The magnitude of the reflection coefficient decreases with increasing $2a/\lambda_0$, i.e. with increasing frequency.

621.374.083.7

S.E.C.R.E. quantized coded "cyclic" telemetering system. A. CLEMENT and J. BRUNO. *Onde Electrique*. **37**, pp. 315-318. March 1957.

In this system data is transmitted in sequential manner, individual information being continuously revised; the recurrence of the revising process depending upon the number of terms in the sequence and upon the band-width available; with a bandwid.h of 620 c/s, every 640 msec for 15 terms, every 1,280 msec for 30 terms, etc. Information is successive.y quantized and coded into binary intelligence. Transmission to the receiver unit is via sub-carrier signals, corresponding to binary code digits. After detection, the six sub-carrier signals provide the display of the figure received on the individual demodulators. These demodulators store up the intelligence until the next information is received and also provide a current proportional to the telemetered data. System accuracy depends upon the number of digits transmitted: 1·5 per cent. for 6 digits, 0·75 per cent. for 7 digits, 0·37 per cent. for 8 digits.

621.374.5

Magnetostriction line. A. ZELEZNIKAR. *Elektrotehniski Vestnik,* **25,** pp. 39-44, January/February 1957.

Magnetostriction lines are used as delay elements in storage loops. For a d.c. pre-magnetization and basing the calculations on the balance of mechanical and electric forces, we can derive symmetrical partial differential equations, observing the line from the transmitting end. A general solution of the wave equation is given. From this equivalent circuits for point $x=0$ are deduced for a line characteristically loaded or unilaterally clamped. The behaviour of an ideal line without losses is also shown. Further formulae determine the resolving time of the system, i.e. time when the proceeding pulse does not exercise essential influence upon the subsequent pulse, the system being bridged by a diode. The resolving time is important for establishing the dimensions of the line, since it determines the maximum iteration frequency of pulses. One of these formulae shows that the cross-section of nickel wire is without essential influence on the resolving time of the system. The delay τ as measured by a short-time meter is 1,450 μsec.

621.376.3

Protection against parasitic frequencies in indirect type frequency standards. R. JAEGER. *Onde Electrique,* **37,** 306-314, March 1957.

Means of protecting against parasitic frequency which are particular to indirect type frequency standards, are discussed. A formula which permits the calculation of the frequency excursion from the pilot frequencies following on a parasitic signal affecting the discriminator, is established and discussed. This formula demonstrates the part played by each of the three factors: frequency plan, synchronism, stability of servo. The limits imposed upon the frequency modulation system are analysed. An example of the determination of the elements is given.

621.385.1

Ceramic envelope electronic tubes. C. OGER. *Onde Electrique,* **37,** pp. 265-269, March 1957.

Ceramic envelopes become a necessity for high power electronic tubes, and are also necessary for medium and low power tubes when very close manufacturing tolerances, or high temperature operation are required. The problems presented to those working in the field of ceramics divide into four categories—the search for a material which will conform with the technical requirements of the tube, the production of this material, the metallization of the surface before bending, development of the metal associated with the ceramic member. The use of ceramics, which has been somewhat delayed in France, opens up the possibility of very high powers in the ultra-short-wave band, and also the automatic manufacture of receiving tubes.

621.385.1.029.6

Development of the strophotron. H. HAGGBLOM and S. TOMNER. *Ericsson Technics,* **12,** No. 2, pp. 165-184, 1956.

The strophotron is an oscillator capable of being tuned electronically by changing the potential of either one of two electrodes with respect to the cathode, and consequently well adapted for frequency modulation. Experimental results from different tubes in the frequency range 1,000–5,000 Mc/s will be reported. A typical value of the electronic tuning range between half-power points is of the order of $\pm$ 2 per cent. The reported work has been concentrated to tubes with power output 1–5 W. Efficiencies of 30 per cent. at 1,000 Mc/s and 10 per cent. at 5,000 Mc/s have been achieved. A small signal theory based on the assumption of current density modulation through a sorting out process will be discussed. Possible power output, efficiency and electronic tuning range are calculated according to this theory.

621.395.625.3

Noise in magnetic recording tapes. P. A. MANN. *Archiv der Elektrischen Ubertragung,* **11,** pp. 97-100, March 1957.

The active layer of magnetic recording tape consists of a random arrangement of grains of magnetic material which in turn are established from a number of Weiss' domains of different orientations. When such tape passes by a reproducing head, a noise voltage appears at its terminals since every Weiss domain gives rise to a voltage pulse. Under the assumption of a fully statistical arrangement of the magnetic domains in respect of position and orientation a formula is derived for the squared voltage per cycle that exhibits a maximum at a certain frequency and disappears for zero frequency and infinite frequency. The noise is proportional to the magnetic moment, i.e. the mean volume of Weiss' domains. This explains why the noise of magnetic tape on which a recording was made and erased again exceeds that of a tape never used before, since the demagnetization applied in the erasing procedure cannot subdivide Weiss' domains to below a certain minimum size.

621.396.812.5

Sporadic ionization of the E-region in middle latitudes. W. BECKER. *Archiv der Elektrischen Ubertragung,* **11,** pp. 101-104, March 1957.

It is shown that the concept "sporadic ionization of the E-region" can cover all ionization steps between 60 and 180 km height that are sporadic in time and hence not due to normal ultra-violet solar irradiation. By reference to observed data it is shown that these "Es layers" evidently are independent of each other and accordingly different mechanisms must be assumed for them. A refined classification of these Es lavers by type and height of layer is proposed. The observations of the Lindau station suggest that in medium latitudes at least five different causes can be assumed for such sporadic ionization observed in the E-region.