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 21 JUNE 1961 NUMBER 6

A CHALLENGE

OUTSTANDING in a year of intense Institution activity was the Dinner held in London on 8th June. Such functions have an increasing attraction and interest perhaps because they are not annual events in the Institution's calendar. This last Dinner provided opportunity not only for social pleasure, but for the propagation of new ideas, which in turn provoked thought on the challenge which the radio engineer faces in the future.

Such a spirit of enterprise was initiated by the President of the Royal Society who spoke of the need for ". . . developing new observations for practical ends." Sir Howard Florey also challenged engineers to think more of technical training from school age upwards; his emphasis on enjoying achievement and work introduced moral philosophy into this material age.

It was also an historical occasion for the Institution inasmuch as for the first time a Past President had again been elected President. Admiral of the Fleet the Earl Mountbatten of Burma, K.G., marked the occasion by emphasizing in his Address the need for increased *engineering* enterprise. The action he proposed in the development of communication techniques makes his Address a pattern of challenge to radio engineers for some years ahead. The Institution's President also referred both to the responsibility and the challenge set by the members themselves in petitioning for the grant of a Royal Charter.

The *background* to that Petition was reflected by Mr. W. E. Miller. With the experience of thirty years of Institution membership, he welcomed guests by emphasizing that the Institution's constant purpose had been to encourage co-operation between engineers throughout the world and especially within the Commonwealth.

The evening's programme was arranged against the theme of the Institution's 7th Convention, shortly to open in the University of Oxford. The purpose of the Convention led the High Commissioner for Canada, The Hon. George Drew, to remind members that the development of radio science had happened almost within a single lifetime and that it was ". . . a challenge to everyone of us to work together for the common good of all."

In this purpose all engineers have an important part to play but members of the Institution now have particular reason, as well as opportunity, to prove their ability to meet the challenge of the space age.

G. D. C.

The Institution's 1961 Dinner

The Institution's Dinner at the Savoy Hotel on Thursday, 8th June, was the first to be held in London since 1958. Nearly 400 members and guests assembled under the Presidency of Admiral of the Fleet the Earl Mountbatten of Burma, K.G., the 16th President of the Institution.

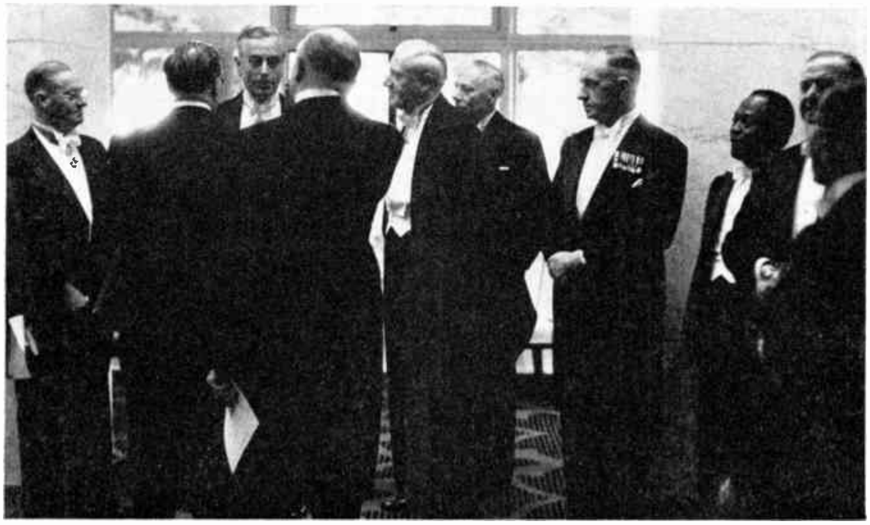
Prior to the Dinner the Council of the Institution had forwarded a message of loyal greetings to Her Majesty The Queen, and the President read the following reply from Her Majesty:

> "Please convey to the members of the British Institution of Radio Engineers and their guests at dinner this evening, my sincere thanks for their loyal greetings which, as Patron of the Institution, I deeply value."

Messages of good wishes received from all the Commonwealth Sections and from representatives oversea were read by Mr. W. E. Miller, M.A., a Past President of the Institution, before he proposed the toast of the Guests.

During the Reception and Dinner, Lieutenant-Colonel F. Vivian Dunn, C.V.O., O.B.E., F.R.A.M., R.M., Principal Director of Music, Royal Marines, conducted the Orchestra of the Royal Marines School of Music.

Members and guests widely represented all fields of science, engineering, education, Government departments, the Services and the radio and electronics industry.



The reception over, a group of guests and officers in discussion with the President.

In addition to the list of guests published in the May *Journal* (p. 388), the Officers of the Institution were also pleased to welcome:

Allen, Sir George, C.B.E. (*Secretary, British Association for the Advancement of Science*)
Asafu-Adjaye, His Excellency The Honourable Sir Edward (*High Commissioner for Ghana*)
Cockburn, Sir Robert, K.B.E., C.B. (*Chief Scientist, Ministry of Aviation*)
Cocking, W. T. (*Editor, "Electronic Technology"*)
Conway, A., B.SC.(ENG.) (*Editor, "Control"*)
Croft, Dr. A. J. (*Administrator, Clarendon Laboratory, Oxford*)
Daeniker, His Excellency The Honourable Armin (*Swiss Ambassador*)
Devereux, F. L., B.SC. (*Editor, "Wireless World"*)
Drew, His Excellency The Honourable George A., Q.C., LL.D. (*High Commissioner for Canada*)
English, C. R. (*Chief Inspector for Technical Education, Ministry of Education*)
Eriks, S. S., O.B.E. (*President, Electronic Valve and Semi-Conductors Manufacturers' Association*)
German, Sir Ronald, C.M.G. (*Director-General, Post Office*)
Haviland, D. L., C.B. (*Deputy Secretary (C), Ministry of Aviation*)
Humbert, Commandant J. (*Air Attaché, French Embassy*)
Humphreys, O. W., C.B.E., B.SC. (*Chairman, Radio Research Board, D.S.I.R.*)
Kaul, T. N. (*Deputy High Commissioner for India*)
Marshall, C. A. (*Editor, "British Communications and Electronics"*)
Phillips, G. C., O.B.E. (*Chief Telecommunications Engineer, Ministry of Aviation*)
Powell, Lt.-Commander C. (*Secretary, Parliamentary and Scientific Committee*)
Reheem, C. K. (*Scientific Liaison Officer for Pakistan*)
Ridgeway, J. W., O.B.E. (*Chairman, Radio Industry Council*)
Stewart, Air Vice Marshal, C. M., C.B.E. (*Director-General Signals, Air Ministry*)
Stockdale, Major-General R. B., O.B.E. (*Headquarters, Technical Group, Royal Electrical and Mechanical Engineers*)
van Nispen, S. C. (*First Secretary, The Royal Netherlands Embassy*)
West, W. R. (*Secretary, British Radio Valve Manufacturers' Association*)

The Importance of Technology

The Speech of Sir Howard Florey, P.R.S., M.D., M.A., Ph.D., F.R.C.P.† in proposing the toast of the Institution.

It is the custom of the Royal Society to choose Presidents alternately from the physical and biological sciences. Tonight you have done me the honour of asking me to propose the Toast of your Institution, but in this you've taken a grave risk, for being a biologist I know next to nothing about electrons and radio waves and naturally I feel somewhat apprehensive at addressing a body of people who are, at least in a technical sense, the direct descendants of Jupiter. He was, as far as I can see, the first large-scale electronics engineer and he must have derived immense pleasure from fiddling with his electrons and hurling his thunderbolts at those who displeased him, though I doubt whether he knew that he was generating electro-magnetic waves at the same time. I suppose that few of you now actually sit on clouds to manipulate your electrons and to generate your radio waves, but I hope that some of you at least have your heads in them. For those in this happy position often produce the most revolutionary thoughts and devices and the success of all your work depends on discovery and the appropriate development of discoveries.

There are welcome signs that the Government is aware that technology has been grossly neglected for many years and great efforts are now being made to enlarge training facilities by expanding the established universities, by founding new ones, by encouraging colleges of advanced technology and in general by trying to improve the technological training of the substantial number of people who in these days are involved in such complicated industrial fields as you cultivate. Nevertheless some scientists fear that the country is not yet applying sufficient of its resources to discovering new phenomena and to developing new observations for practical ends.

It is true that physics, the study of electrons and the sciences in which your institution is particularly interested, are having substantial amounts of the country's resources devoted to them. But are they adequate? For this country is now much more dependent upon its wits for its living that it was even a relatively short time ago and we cannot afford to fail to utilize all our mental resources.

Do we give all our young people the best possible chance of devoting their talents to what they are capable of doing? In other words, is the school system adequate to feed into establishments for higher education an adequate number of intelligent people and is the training in craftsmanship the best possible? There are reasons for supposing that this is not the case, but it's encouraging to see that these matters are now being taken very seriously and perhaps improvements can be expected in the near future. It is perhaps unnecessary to mention such things as these to a great professional association such as yours, for you devote much of your energies to furthering education and technical training—indeed the first chapter of the book on your Institution deals with these matters. Your *Journal*, your conferences, your Clerk Maxwell Lectures are all devised to diffuse information not only to your members in this country but to your affiliates in the Commonwealth, and the importance of the Commonwealth connection is clearly attested by the presence of so many High Commissioners tonight.

Behind many of your educational activities I am sure is your President. It would be impertinent of me to praise one who has been so pre-eminent in many diverse occupations, but I can congratulate your Institution on being so fortunate as to have him at its head. He has long been interested in telecommunications and he was indeed an editor of the Royal Navy's first manual on this subject compiled in the early 1920's. I am sure that you owe a great deal to him for the way in which the splendid aims of your Institution are maintained and under him you will continue to see that the best interests of your profession are served.

The art and science of communication now dominate many of our activities and there is little reason to suppose that your inventions will not further alter at least the outward trappings of life. It would be idle for me, as a layman, to particularize the various remarkable developments of recent years. Some catch the public fancy more than others. We have heard, for example, a lot about radar recently, especially from an historical point of view. Whatever the merits of the dispute which has arisen around the men involved in its early application, there can be no question that radar and radio communication now plays an increasingly important role in all forms of navigation.

Then, of course, radar has been much before the public owing to the activities of the Jodrell Bank radio telescope. Some of you may remember that there was considerable difficulty in financing this instrument. I do not know in detail the reasons for this but, whatever they were, it seems to me fantastic that for the sake of a relatively small sum of money those respon-

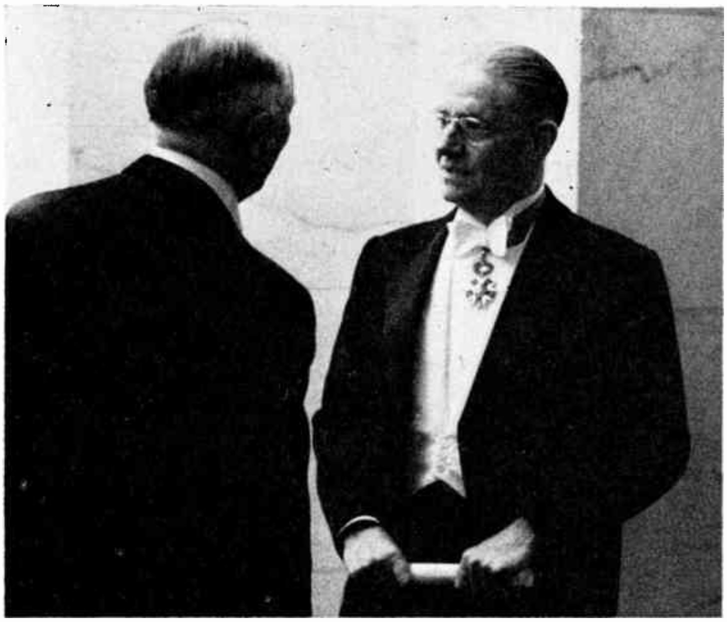
† Professor of Pathology, Oxford University. Sir Howard Florey was elected President of the Royal Society on 29th November 1960.

sible for placing this country in the forefront of this particular application of radio communication should have had to suffer so much in producing the instrument of which the country had reason to be so proud when it was being used to track satellites put up by others.

There must have been few people a decade ago who had sufficient information or imagination to realize that the commercial possibilities of communication systems based on bouncing radio waves off artificial satellites would now be receiving intense attention, but such is the case. Nor, who indeed would have guessed that, as I read in the newspapers not long ago, a group of workers would be bouncing waves off Venus. On musing about the significance of this I tried to recall whether I knew of any paintings in the many galleries throughout the world which would indicate that this practice had been anticipated by Jupiter, but I could think of none—clearly there is much fun in store for radio engineers and maybe there is an idea here for some modern painters.

Electronic devices and miniature communication systems are employed with increasing frequency for research in biology and medicine. A good example is the profound effect that the electron microscope is having on the reassessment of the minute structure of cells and their contents which comprise living matter. There are other interesting possibilities which are only now beginning to be explored. Not long ago at a Royal Society Soirée, I saw a subject who had swallowed a small radio transmitter enclosed in a capsule. From her insides a signal amplified in a most satisfactory way was being transmitted to the outside. I must confess that it passed through my mind that our insides are capable of making a great enough variety of noises by themselves without calling in the aid of electronics to increase them. Nevertheless the possibility of this new form of communication should not be dismissed too lightly for I can well imagine at a future dinner of this Institution each guest will be given a small transmitting capsule with his sherry or his gin or whatever it is he takes before dinner and he would have his own little private receiver behind his ear so that there would be a complete communication system between his insides and his consciousness. Tonight all messages would be of the most felicitous nature, but I can conceive of occasions on which the dinner might not be so admirable so that one might get a private message such as "Don't eat any more of the duck, as it is not doing too well down here". Or "Don't you think you've had enough of that red wine now, you know you've to make a speech in a few minutes". Now that sort of apparatus would be extremely helpful and knowing the potentialities of electronics engineers I think it's quite possible that we shall have such an instrument within a short time. Now if you had given me one of these gadgets tonight, I expect by this time it would be transmitting a loud message to say "It's time you sat down". But before I do so I should like to say that I consider the members of this Institution extraordinarily fortunate people, for you work in a field from which not only practical instruments emerge, but which also yields immense intellectual satisfaction to those who labour in them.

I now have the honour to propose the Toast of The British Institution of Radio Engineers, coupled with the name of your very distinguished President, Lord Mountbatten.



The President of the Royal Society, Sir Howard Florey in discussion with Mr. W. E. Miller, a Past President of the Institution.

The Presidential Address †

of

ADMIRAL OF THE FLEET THE EARL MOUNTBATTEN OF BURMA,

K.G., P.C., G.C.B., G.C.S.I., G.C.I.E., G.C.V.O., D.S.O., LL.D., D.C.L., D.SC.‡

I know that I am speaking for all our members when I say how deeply we appreciate the generous terms Sir Howard Florey has used in proposing the toast of "The British Institution of Radio Engineers". These are all the more valuable to us, coming as they do from the President of the Royal Society himself.

I am sure you all know that we have recently submitted a Petition for the grant of a Royal Charter. This Petition is now before the Privy Council; and you will not, therefore, expect me to speak about it tonight; beyond saying that we all realize that a Royal Charter is an honour and recognition which must be earned, and that all of us have for years been working to this end.

Whatever the outcome may be, I am convinced that we shall continue to do everything in our power to continue the good work on which we have been engaged. It is this conviction which has prompted me to accept for the second time the high office of President of our Institution; and I pledge myself to do all in my power to forward our work.

We are a corporate body under Royal Patronage, who have been dedicated to the promotion of knowledge and new ideas in the field of radio and electronics and who have established our status within the world of engineering.

We must now consolidate our achievements, contributing original thinking, welcoming the new ideas of others and playing our full part in the engineering developments of the future.

The Role of the Engineer

I emphasize the word "engineering": first of all because we are an Institution of Engineers. It is our purpose to encourage the application of engineering technique to scientific discoveries. Whilst this requires our members to understand new scientific ideas and to assist in research, we must not make the mistake of regarding the Institution as a body of research scientists. That function is most admirably served by the Royal Society, whose President and other Fellows we are delighted to have with us this evening.

My second reason for emphasizing the word "engineering" is because I believe that the relationship between science and technology has become slightly confused during these post-war years; and that this has sometimes led to a waste of skilled effort and manpower. Scientific discovery is increasing each year at an exponential rate; but the engineering development of discoveries is still a very arduous task, requiring men of special ability.

The fascination of unravelling the secrets of this universe attracts many men into the scientific field, when in point of fact their bent might well make them invaluable as technologists and engineers. Moreover, there is a crying need today for men who can translate scientific discoveries into practical terms, and help to reduce the time-lag between a new discovery and the general availability of its end-product.

New Computer Techniques

Electronics has been one of the principal tools in the development of automation; and computers, which are playing an increasingly important part in the world, could not have got beyond the stage of an idea without electronic engineering to bring them to life.

Indeed, the principles of the computer were known and adumbrated a hundred years ago; but it is only since the war that advances in electronic technique have made it possible actually to put them into practice. The rise and development of the computer has been phenomenal.

It is only 15 years since I dealt with this newly developing technique in my first presidential address to this Institution; and I remember saying that this technique would soon be able to help in determining the trajectory of projectiles in flight—but I don't think any of us quite realized how very quickly it would become indispensible in guided missile and rocketry development and a major factor in space research.

Nor had I visualized some of the recent uses to which computers have been put. The latest use has been to decipher completely the hieroglyphic manuscripts of the ancient Mayas. The three young Siberian scientists who accomplished this formidable task have just stated that without the use of an

† The First Presidential Address of Admiral of the Fleet the Earl Mountbatten of Burma was published in the *Journal* for December 1946.
‡ Chief of the Defence Staff.

electronic computer this work would have taken centuries, if indeed it could have been done at all.

The technique of the computer has created a demand for a new type of expertise, called "programming"—which is the name now given to the analysis of a problem and its reduction into a language that the computer can interpret. Although some of the donkey work can be carried out by the machine itself, programming is still dependent on human intervention. When correctly programmed a computer will replace a large number of people doing a lot of repetitive mathematical or clerical work; but the few that remain must be very highly qualified if the machine is to be used to full advantage.

There are still laboratory techniques which require engineering development before they can be introduced as reliable components into computers; and there is much engineering to be done to reduce production costs.

There are problems to be solved in the utilization of very large memories; and there is even a need to define English words in precise scientific terms, so that programming techniques can use language intelligible to the machine. The demand now is for computers that have a much greater storage capacity without unduly increasing the size of the machine. In physical design, transistors are easing the problem of bulk; and high speed storage is being improved by the development of very thin magnetic films. There remains, however, the problem of securing easier identification of densely stored data.

Progress on these lines will enable us to improve language translation techniques, and eventually to develop a speech typewriter. For both these purposes we need to develop reliable storage capacity in computers of smaller dimensions. Engineering developments along all these lines are still needed if we are to exploit all the science and engineering that has gone into computer technology during the last few years.

Space Communications

In the field of space communication, which is perhaps the greatest problem we have to face, our Institution has been greatly helped by the initiative taken by the Royal Society in promoting the international exchange of scientific thought on this matter, through its Committee on Space Research, and we are much beholden to them.

As you know, our Institution's Convention at Oxford this year is designed to explore this most important field. The need to plan space communication in terms of electronic engineering is obvious; for without the radio and electronic devices which give it life and link it with the Earth any satellite would, of course, be no more than a piece of inert matter orbiting in space.

It is the delicate instruments that it carries which make it possible for microscopic parcels of energy in various forms to be detected, labelled, recorded, and for the results relayed to observatories on the earth.

The radio waves the satellite transmits are used to study the structure of the ionized layers that surround the earth and to deepen our understanding of the influence of the magnetic fields. Computers then help in analysing all these data; and in this way the range of man's observations has been extended to the edges of the universe and has provided the evidence on which debates on modern cosmology are centred.

1961 will be recalled in history as the year when man made his first ventures into space. But the route had previously been explored by unmanned capsules, armed with electronic instruments and radio transmitters; and it was the data collected by this means that made it possible for the first manned space adventures to be so dramatically successful.

Now that we are entering the Space Age there seems little doubt that unless some quite unforeseen and significant break-through occurs in communication technique the future for reliable global communications lies in the use of satellites as relay stations.

The recent rapid growth of radio in all its forms is bringing about saturation of the ether, since the demand for communication channels continues to grow whilst the available wavelengths are becoming fewer and fewer.

So it seems, at present, as though the provision of suitable communication satellites offers the solution to what will shortly be a very pressing problem. Nor is the problem confined to plain communications.

I expect that the merchant mariner and the civil aviator of the future, like their Royal Navy and Royal Air Force colleagues, will be navigating with the aid of special satellites.

Experiments are being carried out which envisage the satellite issuing instructions to the navigator on demand, which will give the corrections that he has to apply. This will call for an extensive radio system and a large number of calculations made by computers to keep the satellite fed with the constantly changing correction-figures.

Increase in aircraft speeds also calls for very high flying aircraft, and they will depend to an increasing extent on computer analysis of meteorological data. Here the meteorological satellite will be an indispensable aid, provided that the radio engineer can make it an economic proposition by ensuring that the necessary transmitters and electronic equipment will have an endurance of many years.

Our American colleagues successfully launched *Echo I* on 12th August last year, and established communication via this large passive satellite. Although there are obvious attractions to using a *passive* satellite, the signals that can be bounced off it are likely to be so weak that, for really useful communication projects, I believe we will be forced to *active* satellite repeaters, certainly in the case of navigational and meteorological satellites.

This presents us with an immediate electronic requirement to provide reliability; because the satellite must constitute an "unattended repeater" in the fullest sense of the word.

A number of active satellite repeater projects have been formulated, varying from multi-satellites in random orbits to two or three single "stationary" satellites. By "stationary" we infer a 24 hour equatorial orbit co-sensed with the earth's rotation, and this involves orbiting at a height of 22,300 miles. None of these projects would even begin to come within the realms of practicability, if it were not for the new electronic techniques.

All the latest developments in global communications are a particular challenge to British radio engineers; as they are a logical extension of the great work we have done on trans-oceanic cable-links, and other forms of world-wide communications, in the past.

It is immediately clear, for instance, that one of the first challenges will be the colossal premium on miniaturization, in terms of pounds sterling per pound mass for a satellite such as I have just described, particularly as regards launching costs. The economic aspect of developing and implementing our requirements in Space Communications will be one of the features of our Conference.

Whatever projects may be recommended as an outcome of this conference, we may be sure that they will imply immense expenditure, not only in effort, but in money; but it is essential that the whole of the communications space-project must be made to make sense as a matter of economics.

Nor is time on our side for we know that ten of the biggest U.S. radio, telephone, and telegraph companies are at this moment discussing a joint commercial venture to develop a satellite communications system. This venture will be sponsored by a new firm "Communications Satellites Incorporated", who plan to have the first active satellite communications service in operation within about three years at an estimated cost of 90 to 100 million pounds.

The first British satellite will be fired into orbit for us by our American colleagues about the end of this year. It will give us a long sought for opportunity to put into practical operation much of the work done in Great Britain in recent years.

We all recognize that there is still a long way to go, but I am sure this first launching of our own British satellite will encourage us, Radio Engineers in particular, to renew our efforts.

Communications Networks

In my present job I am naturally interested in the way the fighting services develop in the future; and in view of my old specialization in communications I am, of course, particularly interested in the possibility of integrating long-distance communications networks, since it seems obvious that no single defence service—or civil department, for that matter—can really afford to put its own satellites into orbit.

Each of the three Defence Services at present operates its own long distance world-wide communications network; and I am one of those who believe it would be better to combine these three separate networks, thereby saving both money and manpower, but this is a development which can only be logically introduced as the next generation of equipment comes to be designed and put into service. In any case, the Services have their own economic problems caused by changes in the design of communications equipment which are, of course, a continuing process.

It is becoming increasingly difficult for each Service to afford changing its own equipment frequently, however desirable this may be in order to keep it up to date. In the field of wireless telegraphy, for example, great strides are being made in the development of techniques for automatic switching; but the equipment is costed in hundreds of thousands or even millions of pounds, and three such separate switching centres (one for each Service) in, say, Singapore, would obviously be uneconomic from all points of view.

So it seems essential that communications users should get together and pool their resources; since not one of them will be able to "go it alone" in the face of the technological developments which are with us now, and which are growing at an ever increasing rate. I am glad to tell you that all interests, both Service and civilian, are represented at the top official level, in investigations into the possibility of all governmental communications sharing in future expensive developments, such as the use of communications satellites. But there is a limit to the extent to which it is reasonable to save manpower.

We do not want to be like the work-study engineer, who criticized an orchestral performance by saying that it seemed unnecessary duplication that all the first violins played identical notes, since one man could have done the job and the additional volume of sound could have been provided by electronic amplification. I, for one, would never agree to saving manpower by integrating the tactical or forward communications of each service.

I believe that the Navy must still have its own links to enable Naval shore authorities to communicate with the ships in their areas; the Land Force Commanders must have their own links to their formations in the Field; and the Air Force Commanders must have their own links to the aircraft flying under their command.

Education for the Space Age

Everything I have said tonight seems to point one over-riding moral—which is, that in the vast and rapidly expanding field of electronics the duty of all of us who are engaged in the science is to keep in touch with all the gigantic strides that are being made.

I myself qualified as a Naval Communications specialist in 1925—the very year in which our Institution was founded. In the intervening years the accelerating pace of development in radio and electronic techniques and applications has made it virtually impossible for anyone to be the "compleat" radio—or electronic engineer; it has certainly made it difficult for old hands like myself to keep abreast.

It has become of vital importance that we should examine the basic education and vocational training of the younger generation of radio engineers, into whose hands much of the success of future production, operation and maintenance techniques will be entrusted. Our Universities and Technical Colleges have already shown their willingness and ability to help in training the men who are so urgently needed to meet the challenge of the Space Age.

In the Institution, we are proud that my predecessor as President, Professor Zepler, was appointed to the Chair of Electronics in the University of Southampton in 1949—the first Chair of its kind in the country.

Since I have links with that University, I am specially delighted to hear that they have now added to their Faculty a Department of Astronautics in which Electronics is well featured.

The policy of our Institution, with which I was in complete agreement, has been to create specialist groups, which give the maximum opportunity to radio engineers for meeting and working with, colleagues with specialized interests similar to their own. Although it may invite criticism as establishing specialists within a specialized body, the Group system has real advantages. Under the umbrella of general meetings of the Institution, we still preserve the means to encourage cross-fertilization of ideas between those whose work is primarily in the fields of radar, television, computer, or instrument technology, for example.

Group activities also help to direct our attention to new applications of radio or electronic technology. This has been particularly evident in the field of electronics in medicine and biology, as the President of the Royal Society has reminded us.

The Electro-acoustics Group of our Institution is concerned with propagation, conduction and measurement of all kinds of material vibration, and some of their work has been fed back into the work of other groups—for example, their work on hearing aids has been passed to the Medical Electronics Group. We complete the full-cycle when we recall that it is this sort of specialist activity which will do much to help solve the problem of verbal or speech typewriters, which I referred to earlier on.

I have only mentioned this particular Group in order to illustrate that, whatever his speciality, the radio and electronic engineer has full opportunity to play his part in the developments of the future.

For the reasons I have mentioned, the engineer must provide for the maximum efficiency of the instruments he makes, in order not only to ensure the full usefulness of his equipment, but also to reduce its cost. To cope with this vast but vital programme of developments, which will enable us to keep our place in global communications, we will require an ever-increasing number of fully trained radio and electronic engineers.

I doubt whether we are keeping pace with the demands, and although our Institution can do much to forward their technical training we must have an adequate supply of suitably educated young men (and why not women too?) to fill this great demand. This is the challenge which faces our country today.

I have tried, within the compass of an after-dinner speech, to put before you my conception of the way in which radio and electronic engineering should be developing, and the targets it should be setting itself.

But the field is so vast, and the different subjects so specialized, that clearly an organization is needed: an organization which will forward the work of the specialist groups, while co-ordinating their general purpose.

It is precisely this function, so vital to our future, which I conceive it to be the duty of the British Institution of Radio Engineers to fulfil; and I am proud to have had the honour to respond to the toast of so worth-while an Institution.

The Guests

A welcome to the Institution's guests was proposed by Mr. William E. Miller, M.A. (a Past President)

"Members of our Institution have plenty of opportunities to meet each other at our technical meetings and Conventions. It would, however, be very dull if we kept ourselves in a technical vacuum and did not share our social events with friends.

"Tonight we have good reasons for celebration.

"Firstly, because earlier this week, as already you know, our membership unanimously elected Admiral the Earl Mountbatten of Burma as our President for a second term—the first time that we have ever re-elected a Past President. In addition, we are looking forward with some excitement to the Convention we are to hold next month. For both these reasons we are very pleased to welcome tonight many friends of the Institution to share in our pleasure.

"Ten years ago the Institution held a Television Convention in Cambridge and we were very interested to hear a paper on large-screen television presented by an eminent Swiss engineer. Our President has referred tonight to developments in computers, which involve high speed storage and again, I believe, Swiss patents are utilized. The Swiss nation has made notable contributions in a number of technological fields, and as engineers we are honoured and pleased to welcome the Swiss Ambassador, His Excellency Monsieur Armin Daeniker.

"It has always been our policy to encourage co-operation between radio engineers throughout the Commonwealth. It is with particular pleasure, therefore, that we express our appreciation of the presence tonight of the High Commissioners of Canada, Ghana, Malaya, Pakistan, Ceylon and New Zealand. We are also happy to welcome the Deputy High Commissioner for India and also the representative of Rhodesia and Nyasaland and the representatives of the French and Dutch Ambassadors.

"Autonomous sections of the Institution are already established in India, Pakistan, New Zealand and Canada. We are anxious to encourage the establishment of sections in other countries so that members throughout the Commonwealth, and indeed, in many other countries of the world, may enjoy the same opportunities of meetings as we have in Great Britain.

"As an Institution we also believe that it is very necessary to encourage the closest co-operation between radio and electronic engineers, whether they are engaged in industry, in research establishments, broadcasting corporations, or the Services.

"The Royal Navy has played an important part in the development of radio and has always supported our Institution. The closeness of our association with the Senior Service is exemplified by the presence of the First Sea Lord, Sir Caspar John, and his colleagues on the Board of Admiralty.

"During the lifetime of the Institution the demands of the Army and the Royal Air Force for radio equipment naturally have influenced the development of our science. Indeed, the whole history of radar dates back to the initiative originally taken by the Air Ministry. No gathering of this kind would therefore be complete without the presence of Major-General Cole, Air Vice-Marshal Pretty and many of their colleagues from the War Office and the Air Ministry.

"The need for adequate opportunities for the study of radio engineering and the promotion of reliable standards of qualification were two of the main reasons for the founding of the Institution in 1925. For the past thirty-six years the Institution has had many meetings with the interested Government Departments and the education and training branches of the three Services.

"If at times we have been forced to accept a compromise, it has been in the best British tradition, and the friendliness and co-operation that have attended the discussion of matters so important in the world today are evidenced by the presence tonight of representatives of the Ministry of Education.

"Tonight provides us with an opportunity of expressing to them all our thanks for the tolerance with which they accept the Institution's frequent approaches, and especially our appreciation for the way in which they co-operate with us.

"The business of Parliament is largely a matter of endeavouring to influence opinion on any subject that may be under discussion. Many of these subjects today are concerned in some way with scientific and engineering matters, and I sometimes think that the country would be well served by the infiltration of a few more engineers into Parliament. Be that as it may, we of the Institution feel that the Parliamentary and Scientific Committee is an excellent body for providing the much-needed link between scientific bodies and Parliament.

"Incidentally, the work of the British Parliamentary and Scientific Committee is now being emulated not only in other parts of the Commonwealth but also by European countries. Our own association with the Parliamentary and Scientific Committee goes back

20 years, and it is with particular warmth that we welcome Dr. Bennett, the Chairman of the Committee, and Commander Powell, its first and present Secretary, and other Members of Parliament who are with us this evening.

"It would be impossible to think of any other Institution that is held in such high esteem as the Royal Society. It is the father of all such Institutions as our own, and we have much appreciated the honour of having the very busy President of the Royal Society with us. He is supported by a number of Fellows of the Royal Society, including Sir Lawrence Bragg, who gave the third Clerk Maxwell Memorial Lecture to our Institution."

Mr. Miller then referred to the Institution's pleasure in welcoming representatives of learned societies and educational organizations and in asking his audience's indulgence in not referring to them all by name commented:

"Indeed, I am very much in the position of the young student who was asked to name some of the great kings in our history. He covered his inability by answering, 'Sir, they were ALL great!'

"I mentioned earlier that we regarded our work as a contribution toward Commonwealth and international co-operation, an object that must surely be one step nearer achievement as we join together to contribute to this year's convention. This co-operation is evident from the fact that many of the papers to be presented at our Convention will come from the countries of the Commonwealth, the United States, France, and, we hope, the Soviet Union.

"We are particularly pleased that we shall enjoy Canadian participation, because this is the first convention that we have held since the establishment of new Institution Sections in Toronto and Montreal. At the request of our members in Canada, the Secretary of the Institution, Mr. Graham Clifford, is to visit our new Sections in the autumn. You will understand, therefore, that I give a special welcome to the High Commissioner for Canada—Mr. Drew—who is widely appreciated in London for the considerable work he has done, not only for his own country but in promoting good Commonwealth relations.

"On behalf of the Officers and Members of the Institution, I again thank all our guests for their attendance, and give to our members the toast of Our Guests, coupled with the name of the High Commissioner of Canada."

Admiral of the Fleet The Earl Mountbatten of Burma, K.G., was the 9th and is now the 16th President of the Institution. His comment about this amuses Mr. L. H. Bedford, C.B.E., the 10th President, and Mr. W. E. Miller, M.A., who was the 12th President of the Institution.

Cooperation within the Commonwealth

The reply made by His Excellency the Hon. George A. Drew, Q.C., LL.D., High Commissioner for Canada in the United Kingdom, to the Toast of the Guests.

The Toast of the Guests offers convincing proof that, in addition to his other attainments, Mr. Miller is eminently qualified for the Diplomatic Service! I know that all your guests would wish me to express our thanks for the privilege of being with you this evening and to say how grateful we are for the warmth of your reception. We also thank Sir Howard Florey for his very clear interpretation of the role that your Institution plays in promoting wider international understanding of the challenging possibilities of the application of electronics to medical and biological research.

We have all been stirred by the appeal of Lord Mountbatten for more highly trained engineers to translate scientific discoveries into practical results and reduce the time lag between invention and production. I have had the opportunity to see how effectively he converts theory into practice, and I am confident that under his inspired and inspiring leadership the new objectives now set before the Institution will be attained. The vast strides in the science of engineering, and particularly space research, since he was your President just after the war, emphasize the speed with which we are moving in this vitally important field. I congratulate you on your choice of President, and I extend to him every good wish in the new task which he has undertaken with such typical enthusiasm and energy.

I have no doubt that the reason you have done me the honour of asking that I speak on behalf of the guests tonight is that this Institution has been so active in building up wider scientific contacts within the Commonwealth. While there are representatives here from many other countries, I know that your direct contacts overseas have been more and more with members of the Commonwealth. I hope all of you will understand that this is the reason why I shall devote my remarks mainly to events that have already happened within the Commonwealth, merely as an indication of what may be achieved by a free exchange of ideas and the more effective co-ordination of research and production in the very costly field of new radio techniques and space research. What I have to say will, I think, apply with equal force to every other country represented here, but quite frankly I am going to say something of the immense achievements of the countries of the Commonwealth, and primarily Britain, at a time when there seems to be some tendency to underestimate the role that we have played.

History is no mere echo of a dead past when it points the way to the future. I think this is an appropriate occasion to recall that it was in 1867 that Clerk Maxwell started the dramatic chain of events which step by step has brought to us the amazing radio communications that we have today. It was in 1867, the year Canada became a nation, that he read a paper to the Royal Society, in which he put forward the theories which form the basis of wireless as it was then termed.

The beginning of radio in Russia, which has played such an important part in their extremely successful space research, dates from the last decade of the nineteenth century as a result of the work of Popov and Rybkin at the same time as the Royal Navy was experimenting with this new method of communication.

It was Sir William Watson who as far back as 1747 started electrical transmission of messages when he demonstrated that words could be carried electrically over a considerable length of wire. It was a Scotsman, born and educated in Edinburgh, who invented the telephone in Boston in 1876, but in the meantime he had been living in Canada and returned to spend the latter part of his life there, during which he made dramatic contributions to the newer methods of transportation and communication. It was in an aeroplane made by Bell in Nova Scotia that McCurdy made the first controlled flight in the British Empire at Badeck in February 1909. It is perhaps also a matter of interest that during the First World War Bell designed and built the first hydrofoils, similar in design to those now being made in Italy and Japan, and incidentally it was Bell who first used that name.

As further evidence of the close association in this progressive field of research between Canada and Britain, the first radio signal ever sent across an ocean was that sent by Marconi from Cornwall to St. John's, Newfoundland, in December 1901.

We have come a long way in a comparatively short time from the earliest telegraph, telephone and radio messages to the modern trans-Atlantic cables which we have today. Again I hope I may be forgiven for recalling that the first trans-Atlantic telephone cable was laid as recently as 1956 jointly by the British Post Office, the Canadian Overseas Telecommunications Corporation, and the American Telephone and Telegraph Company. By 1964 there will be a British Commonwealth cable service around the world. The first link in this world-wide system will be laid this year between Newfoundland and Scotland and will handle 80 simultaneous voice channels over the same cable. My reason for mentioning some of these incidents, perhaps not on the surface immediately connected with space research, is to show what can be done by co-operation and what has been achieved primarily by British research in this field.

Certainly in thinking of space research, radar comes immediately to our mind. This again was a British invention which had a profound effect on the outcome of the last war. In fact our Russian scientific friends, who have made such remarkable advances in space research, received their first radar equipment during the war, when the Western allies supplied them so generously with military equipment of all kinds which made it possible for them to survive.

What we must recognize, however, is that they have been able to rationalize production and allocate tasks for a very large population which has a tremendous pent-up consumer demand.

The lesson for us, and I believe there is a lesson, is that with our much greater combined technical skill we should gain the full advantage of our long years of experience by working together so that we may, on a voluntary basis, gain the most from our experience and opportunities, by the integration of research efforts and the allocation of production, in a way that will produce the best results from the resources at our command. Now we are entering a new phase of research which before long will give us world-wide satellite communication. In this field Australia has made an immense contribution by the work done at the Woomera Ranges. What was accomplished there has had much to do with putting earlier satellites into orbit.

It now seems likely that there will be world-wide satellite communication in effective operation by 1970. Such a system will have a fantastic capacity for the simultaneous handling of television, radio and voice communication. Soon there will be no secrecy for any of us. It may have the advantage, however, of letting everyone know how everyone else is living and what they are doing.

The Institution's Canadian Sections are discussed by the President and the Hon. George Drew with Mr. G. D. Clifford who will visit Canada in the autumn.

Surely what has happened almost within a single lifetime is a challenge to everyone of us to work together for the common good of all. Within this Commonwealth are vast land areas and limitless resources combined with great accumulated skill. Working together there is nothing we cannot achieve. That is the challenge and that is the hope placed before all of us. In expressing our thanks I am sure I may say, on behalf of all the guests here tonight, that nobody could be better fitted to undertake this task than your new President who has pictured so clearly the bright new horizons now opening out before our eyes.

Brit.I.R.E. PAPERS ON RADIO TECHNIQUES FOR SPACE RESEARCH

During the last few years a number of papers have been published in the *Journal* on techniques which are being employed in space research and communications. The following list, which will be of particular interest to members attending the Convention, includes some of the more significant papers on space techniques, telemetry and propagation.

A Rocket-Borne Magnetometer. K. Burrows

This paper describes a proton precession magnetometer designed to measure the earth's magnetic field. The equipment was developed for firing in a *Skylark* research rocket as part of the Royal Society's upper atmosphere research programme in connection with the I.G.Y.

(Vol. 19, No. 12, pp. 769–776, December 1959)

System Design Criteria for Space Television. A. J. Viterbi

A narrow band television system for relaying to earth images of the planets is described. Bandwidth compression is achieved by storing the video signal on magnetic tape or photographic film and transmitting at a reduced information rate. An f.m.–p.m. telemetry system is utilized.

(Vol. 19, No. 9, pp. 561–570, September 1959)

A Method for Interpreting the Doppler Curves of Artificial Satellites. G. Boudouris

A graphical-analytical method designed for determining the point of inflexion and the maximum slope of experimental Doppler curves and hence the time of passage of a satellite at its minimum approach distance is described.

(Vol. 20, No. 12, pp. 933–935, December 1960)

Telemetry Aerials for High-Speed Test Vehicles. R. E. Beagles

The problems peculiar to propagation of telemetry signals to and from high speed test vehicles are discussed and typical external and suppressed aerials described.

(Vol. 20, No. 8, pp. 497–504, August 1958)

A Comparison of Telemetry Systems. A. Cowie

This short introductory contribution to a Symposium on "Radio Telemetry" discusses the factors influencing the choice of modulation systems to be employed and reviews considerations behind the selection of multiplexing arrangements.

(Vol. 19, No. 8, pp. 491–492, August 1959)

Engineering Aspects of Missile Telemetry Equipment—The Airborne Sender for 24-channel Telemetry. W. M. Rae

Design, production and application problems of this a.m.–f.m. system are described.

(Vol. 21, No. 1, pp. 57–67, January 1961)

455 Mc/s Telemetry Ground Equipment. F. F. Thomas

The equipment comprises the receivers, a 16-tube slow speed recorder and a histogram recorder with data processing facilities. The test set for checking and calibrating telemetry sender is also described.

(Vol. 21, No. 1, pp. 69–77, January 1961)

Transportable Ground Receiving and Recording Equipment for 24-channel Telemetry. F. G. Diver

This equipment is developed from that described in the preceding paper (Thomas) and provides a number of improvements to facilitate its operational use.

(Vol. 20, No. 6, pp. 457–464, June 1960)

A Six-channel High-frequency Telemetry System. T. C. R. S. Fowler

This frequency multiplex f.m.–a.m. system permits frequency components from 10 c/s to 10 kc/s to be telemetered simultaneously.

(Vol. 19, No. 8, pp. 493–507, August 1959)

A 600 Mc/s Pulse Telemetry System. G. T. Hardwick

In this pulse position modulation system information is conveyed by variation of the delay time between a wide datum pulse and up to 20 narrow pulses.

(Vol. 20, No. 6, pp. 465–473, June 1960)

An Equipment for Automatically Processing Time Multiplexed Telemetry Data. N. Purnell and T. T. Walters

The output of this data processing equipment is available as an analogue graph on paper film and a digital recorders on punched cards.

(Vol. 21, No. 3, pp. 257–274, March 1961)

Radio Studies during the International Geophysical Year 1957–8. W. J. G. Beynon

The ionospheric investigation to be carried out during the I.G.Y. are discussed including the experiments using rockets and satellites.

(Vol. 18, No. 7, pp. 401–416, July 1958)

Photo-Electric Image Techniques in Astronomy. B. V. Somes-Charlton

In addition to discussing the techniques of the employment of television cameras with astronomical telescopes, a space orbiting astronomical telescope is briefly described.

(Vol. 19, No. 7, pp. 417–435, July 1959)

New Developments in Silicon Photo-voltaic Devices M. B. Prince and M. Wolf

The silicon solar cells described in this paper are of the type used for power generation in artificial earth satellites.

(Vol. 18, No. 10, pp. 583–595, October 1958)

Research in Radio Astronomy.

A brief description of the Mullard Radio Astronomy Observatory at Cambridge.

(Vol. 17, No. 8, pp. 405–406, August 1957)

The 1961 Convention

"RADIO TECHNIQUES AND SPACE RESEARCH"—OXFORD, 5th–9th JULY

Time Table and Programme

Wednesday, 5th July

10.00 a.m.–2.00 p.m. Registration in Christ Church

2.15 p.m. Official opening of Convention in the Clarendon Laboratory

2.30–5.30 p.m. SESSION 1: **Introduction—Radio Techniques in Space Research**

Chairman: I. MADDOCK, O.B.E.

"Satellite Launching Possibilities". D. J. LYONS

"Critical Engineering Factors in the Design and Development of Space Systems". J. M. BRIDGES

"Radio Tracking of Artificial Satellites". DR. B. G. PRESSEY

"The Scientific Uses of Earth Satellites". DR. J. H. BLYTHE

"Some Radio Astronomy Techniques". DR. R. C. JENNISON

7.30–9.00 p.m. Group discussions and technical films

8.00 p.m. A Concert by the Hallé Orchestra in Oxford Town Hall

Thursday, 6th July

9.30–12.30 p.m. SESSION 2: **Satellite Engineering—Systems**

Chairman: L. H. BEDFORD, C.B.E.

"Engineering Aspects of Satellites and their Launching Rockets". G. K. C. PARDOE

"Satellite Launching Methods". C. R. HUME, M.B.E. and R. A. SHUTE

"An Economical and Timely Technique for Conducting Radio Research in Space". J. D. NICOLAIDES

"Telemetry Systems for Space Research". A. E. DENT, W. M. RAE and J. H. WHITE

"Inertial Systems in Space Vehicles". M. A. V. MATTHEWS

2.30–5.30 p.m. SESSION 3: **Satellite Engineering—Components**

Chairman: L. H. BEDFORD, C.B.E.

"The Effect of Environment on Satellite Engineering". DR. R. INNES

"The Reliability of Components in Satellites". G. W. A. DUMMER, M.B.E.

"Techniques in Microminiaturization". D. H. ROBERTS and D. S. CAMPBELL

"Power Supplies for Space Vehicles". K. E. V. WILLIS

"Some Thermal Considerations on the Use of Silicon Solar Cells in Earth Satellites". R. P. HOWSON, D. H. ROBERTS and B. L. H. WILSON

"Solar Cells for Communication Satellites in the Van Allen Belt". DR. F. M. SMITS, K. D. SMITH and W. L. BROWN

7.30–9.00 p.m. Group discussions and technical films

Friday, 7th July

9.30–12.30 p.m. SESSION 6: **Communication Satellites—Engineering**

Chairman: DR. DENIS TAYLOR

"Optimum System Engineering for Satellite Communication Links". W. L. WRIGHT and S. A. W. JOLLIFFE

"Long-Distance Communications via the Moon". P. A. WEBSTER

"Communications at Megamile Ranges". DR. R. C. HANSEN and R. G. STEPHENSON

"Overall System Requirements for Low Noise Performance". C. R. DITCHFIELD

"Some Types of Low Noise Amplifiers". R. HEARN, R. J. BENNETT and B. A. WIND

"The Maser and its Application to Satellite Communication Systems". P. HLAWICZKA

Friday, 7th July (*Cont.*)

2.15–5.30 p.m. SESSION 7: **Communication Satellites—Systems**

Chairman: DR. DENIS TAYLOR

"The Need for Fixed-service Satellite Communications Systems". M. TELFORD and G. A. ISTED

"The Advantage of Attitude Stabilization and Station Keeping in Communication Satellite Orbits". DR. W. F. HILTON and B. STEWART

"A Proposal for an Active Communication Satellite System based on Inclined Elliptic Orbits". B. BUSS and J. R. MILLBURN

"The Quest for Reliable Earth-Space Communications". M. T. HAYES

"Television Communications using Earth Satellite Vehicles". L. F. MATHEWS

"Navigation Satellites with particular reference to Radio Methods of Observation". W. A. JOHNSON

"Some International Aspects of Satellite Communication System Planning". CAPT. C. F. BOOTH, C.B.E.

7.15 p.m. Reception in Tom Quadrangle, Christ Church

7.45 p.m. Convention Banquet in Hall under the Chairmanship of the President

Saturday, 8th July

9.30–12.30 p.m. and 2.30–3.30 p.m. SESSION 4: **Extra-Terrestrial Measurements**

Chairman: I. MADDOCK, O.B.E.

"The Use of Probing Electrodes in the Study of the Ionosphere". DR. R. L. F. BOYD

"Measurements of Solar X-Radiation". DR. K. A. POUNDS

"X-ray Spectrometer for Scout Satellite". J. ACKROYD, R. I. EVANS and P. WALKER

"Cosmic Ray Measurements in the U.K. Scout 1 Satellite". PROFESSOR H. ELLIOT, DR. J. J. QUENBY, A. C. DURNEY and D. W. MAYNE

"Magnetic Measurements in the Ionosphere by Optical Pumping Magnetometer". DR. J. BLAMONT and P. COUFLEAU

"Rocket Measurements of the Upper Ionosphere by a Radio Propagation Technique". DR. S. J. BAUER and J. E. JACKSON

"The Canadian Defence Research Board Topside Sounder Satellite". DR. R. C. LANGILLE and J. C. W. SCOTT

"Ultra-Violet Astronomy from Rockets and Satellites". DR. D. W. O. HEDDLE

"Radio and Photographic Observations of Satellites". DR. D. H. SHINN and N. R. PHELP

3.45–5.30 p.m. SESSION 5: **Techniques in Radio Astronomy**

Chairman: PROFESSOR EMRYS WILLIAMS

"Radio Astronomy from Rockets and Satellites". DR. F. GRAHAM SMITH

"Satellite Techniques for Performing a High Resolution Survey of the Radio Sky at Medium Wavelengths". DR. R. C. JENNISON

"The Australian 210 ft Radio Telescope". DR. E. G. BOWEN and J. P. WILD

"Radar Investigations of the Upper Atmosphere". DR. J. S. GREENHOW and DR. C. D. WATKINS

"Radar Observations of the Planet Venus". L. MALLING and DR. S. GOLOMB

NOTE: *Additional contributions to the Convention will be included in the official programme and the daily "Brit.I.R.E. Convention News."*

Sunday, 9th July

Matins in Christ Church

Convention disperses

ADDITIONAL CONVENTION PAPERS

The April and May *Journals* included synopses of 21 papers which will be presented during the Convention. Further papers accepted are summarized below, but there will be additions. Members are reminded that the availability of reprints will be indicated under "Papers Accepted for Future Publication" (see p. xxxiv of this issue).

The Effect of Environment on Satellite Engineering

SESSION 3

R. INNES, B.SC., PH.D. (*Marconi's Research Laboratory.*)

The radiation environments encountered by a satellite according to its orbit are summarized. The breakdown of high energy radiation on meeting matter is described, indicating the mechanisms and nature of damage caused to equipment. An argument is presented to show that screening against hard radiation is inefficient and uneconomical. Estimates are given of the life of a transistor—the most susceptible to damage of normal electronic components—in a satellite traversing the first Van Allen belt. Means of protection against the erosive low energy radiation are discussed and requirements to meet launching conditions are mentioned.

Radio Astronomy from Rockets and Satellites

SESSION 5

F. GRAHAM SMITH, PH.D. (*Mullard Radio Astronomy Observatory, University of Cambridge.*)

It is not possible to mount an aerial system with a high gain in a rocket or satellite exploring the strength of radiation above the ionosphere, as the wavelength is too large. A measurement of the average emission from a hemisphere, using a simple dipole can, however, be related to similar measurements at higher frequencies. This experiment will be incorporated in the second U.K. *Scout* satellite. A successful test of the apparatus has already been made by a flight in a *Black Knight* rocket, which gave a tentative value of background intensity at frequencies in the range 0·75 to 3 Mc/s. The aerial system for U.K. *Scout 2* will be a long wire dipole, and this is to be tested in a *Skylark* rocket experiment.

Some Types of Low Noise Amplifiers

SESSION 6

R. HEARN, B.SC., R. J. BENNETT AND B. A. WIND, B.SC. (*The Plessey Company Limited, Ilford.*)

The need for amplifiers with optimum noise performance will be discussed with special reference to space communications. The paper will then deal with the application of conventional triodes, parametric amplifiers (quadripole and varactor), tunnel diodes and transistors, to the problem of obtaining a low noise receiver in the frequency range 100–1000 Mc/s. A critical survey will be given for thermionic triodes, quadrupole, and tunnel diode amplifiers, indicating advantages and disadvantages of these devices.

Experimental results, with a brief outline of the underlying theory, will be given for a varactor amplifier used as a preamplifier to a u.h.f. receiver. Results will also be given for a transistor used as a self-oscillating mixer which shows conversion gain and a good noise performance. Consideration will then be given to improving the performance of the devices with particular reference to increasing the bandwidth of varactor and tunnel diode amplifiers.

The Maser and its Application to Satellite Communications Systems

SESSION 6

P. HLAWICZKA, B.SC.(ENG.), B.SC. (*Marconi's Research Laboratory.*)

The general principles of operation and construction of the maser are reviewed, with emphasis on its ability to give low noise amplification. The performance of the travelling wave maser, achieved and expected, in respect of gain, bandwidth, noise factor and frequency are examined, and the relation of the last two to the microwave "space window" considered. The practical requirements of an installation, including accessory equipment, are described, and consideration given to the problems of attaching it to a communication system. Some consideration of possible cost is included. Finally, experimental data will be given on a travelling wave maser constructed.

The Need for Fixed-Service Satellite Communications System

SESSION 7

M. TELFORD, B.SC.(ENG.) AND G. A. ISTED. (*Marconi's Wireless Telegraph Co. Ltd.*)

Long distance radio communication has developed in a somewhat haphazard fashion in the past, and there is little doubt that the world is now faced with the problem on how to expand this facility in a logical manner. The degree of expansion required will be considered in the paper by reference to the available statistics on long distance telecommunications and by analogy with other forms of communication. The revenue from a hypothetical satellite communication system is discussed with reference to capital outlay and annual cost. A plea is made for the highest degree of international collaboration in the planning and operating of such systems.

U.D.C. 621.382 : 620.169

Life Characteristics of Some Typical Semi-Conductor Devices

By

R. BREWER †

AND

D. J. E. RICHARDS, B.Sc.‡

Presented at the Symposium on New Components held in London on 26th–27th October 1960.

Summary: Evidence from life tests of transistors is presented as an indication to equipment designers of the changes in electrical characteristics which may occur during the use of these devices. The effects of physical tests such as those for storage at high and low temperatures, high humidity, vibration and shock are also reviewed, and an indication is given of the incidence of inoperative failures. Operational reports of some equipments using large numbers of semi-conductor devices are compared with laboratory life test evidence.

1. Introduction

In the evolution of any new product the assessment of its performance under normal conditions of usage inevitably comes at the end of the long line of development from inception to commercial production. Many types of semi-conductor diodes and transistors may not be regarded as "new" components, but it is only now, when some of them have been in production for a few years, that their life pattern is beginning to be established. The information given in this paper is therefore the result of early attempts to establish suitable techniques for the assessment, interpretation and presentation of information about the reliability of semi-conductor devices, the findings being based on life tests and field reports of some typical devices that have been in production for 3 or 4 years.

To the user, information about the life characteristics of devices in large-scale production is of greater interest than that for devices which are still in the development stage. Although the latter may show promise of high orders of life performance, ultimately it is the variations in life characteristics due to small changes in production processes which largely determine the behaviour under normal working conditions. The interval between the development and final production stage for some types of devices may be lengthy and it follows therefore that we may have to wait until the late 1960's to obtain reliability information about components made under full-scale production conditions which are now in the development stage.

† The General Electric Company Limited, Central Research Laboratories, Hirst Research Centre, Wembley, England.

‡ Formerly with the Research Laboratories of The General Electric Company Limited, now with National Union of Manufacturers' Advisory Service Limited.

2. An Outline of the Life Characteristics of some Typical Semi-conductor Devices

From the equipment designers' viewpoint the main pattern of transistor and semi-conductor diode life characteristics is now sufficiently well known as not to require elaboration. It is basically similar to that of many other components in which "failures" may be grouped under two main headings:

(1) Complete inoperative failure (sometimes known as catastrophic failure).

(2) Deterioration of one or more electrical characteristics (sometimes known as characteristic or degradation failure).

Dealing with the latter first it has been found from life tests under electrical operating conditions that in germanium transistors some fall in current-gain may occur, the extent of which is dependent upon the junction temperature. Collector cut-off current may

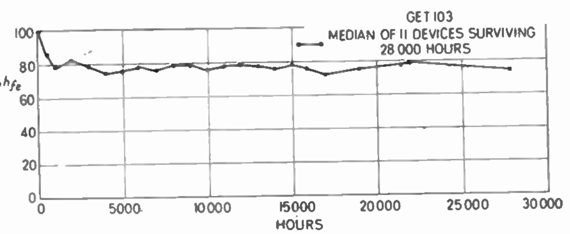


Fig. 1. GET103. Median h_{fe} in electrical life test at $T_j = 65°$ C.

increase slightly during life but for many types of device both of these changes are more rapid in the first few thousand hours of life than at later periods. Indeed some life tests have shown that after a few thousand hours current-gain may become almost stable, as shown in Fig. 1. The evidence in this figure is based on the results of 11 germanium transistors which were placed on life test in 1957. The devices are operating in an ambient temperature of 45° C with an electrical dissipation to give a junction temperature of 65° C, which was the maximum rating of this device at the time of manufacture. The measurement of electrical characteristics is carried out at 25° C and the record of small-signal current-gain (h_{fe}) shown in Fig. 1 is of measurements made at this temperature. It will be seen that after falling about 30% by 7000 hours, h_{fe} has remained substantially stable to nearly 30 000 hours, which is as far as the test has run.

As the result of improvements in the manufacturing techniques, the maximum junction temperature of this particular type of transistor was increased to 85° C and devices now being made have approximately the same order of h_{fe} degradation at 85° C as the older version had at 65° C.

No marked change occurs in the spread of h_{fe} during life, as can be seen in Fig. 2. This shows the distribution of h_{fe} in a random sample of GET103 transistors before and after a 500-hour electrical life test at a junction temperature of 85° C. The distribution is plotted on arithmetic probability paper so as to show its approximation to a Gaussian distribution, and to facilitate comparison between the initial and the 500-hour measurements. The latter are shown by the crosses. It should be noted that the distributions are derived from histograms and that initial and 500-hour points at any given percentage level are not necessarily composed of the same specimens. The effects of operation at different junction temperatures on the principal characteristics h_{fe} and I_{cb0} are shown by Figs. 3 and 4 respectively. These show the results of life tests of the same type of germanium transistor operated at junction temperatures of 65° C and 85° C. Again it will be noted that most of the fall in h_{fe} occurs in the early part of life but that the 20° C reduction in junction temperature raises the level at which h_{fe} begins to stabilize by about 30%.

It should be emphasized that the foregoing evidence has been taken to illustrate the type of results obtained from routine life tests of devices in large-scale production. Other types of transistors may show considerable differences in life behaviour from those given, and one of the difficulties confronting the equipment designer is that of obtaining a general statement which is valid for transistors as a whole. There is little doubt that at the present stage of development any such statement which is intended to cover a wide range of devices, would necessarily have to be so broad as to be often unrealistic for the

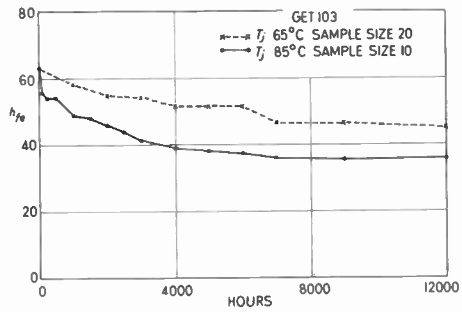


Fig. 3. GET103. Median h_{fe} in electrical life tests at T_j = 65°C and 85° C. Ambient temperature 55° C.

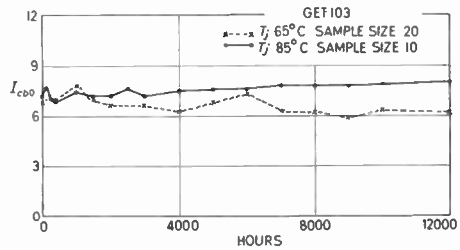


Fig. 4. GET103. Median I_{cbo} (at 25° C) in electrical life tests at T_j = 65° C and 85° C.

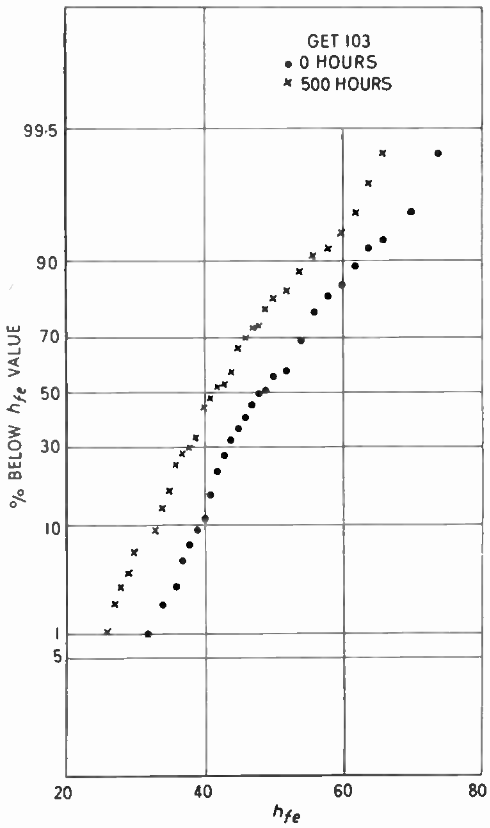


Fig. 2. GET103. Distribution of h_{fe} before and after 500-hour electrical life test at T_j = 85° C. (Arithmetic probability scale.)

circumstances obtaining in any particular application. Equipment designers should therefore endeavour to obtain an estimation of the life behaviour of the particular devices that they intend to use and make the appropriate allowance for changes likely to occur during life in the principal characteristics.

3. Evidence from the Services' CV7000 Specifications

Valuable additional evidence about the reliability of some semi-conductor devices has been obtained from the various tests laid down in the Services' CV7000 series of specifications. These specifications are issued on behalf of the three Armed Services, the G.P.O. and the U.K. Atomic Energy Authority for the purpose of procuring semi-conductor devices of known characteristics and reliability. The latter aspect is covered by two series of tests referred to as Groups E and F. The former includes mechanical and environmental tests such as those for vibration-fatigue, temperature and humidity cycling, shock, lead fragility and soldering. The Group F tests consist of electrical operation to 1000 hours, and high- and low-temperature storage for 150 hours. Devices offered to the CV7000 specifications are grouped into weekly lots, each lot being held in bond until the completion of all the specified tests. The whole procedure is essentially the same as that evolved some years ago for the CV4000 series of "reliable" valves. It is of interest to note, however, that while the introduction of test clauses calling for the assessment of reliability throughout production was a late development in CV valve specifications, this requirement has virtually been standardized in the CV specifications for all semi-conductor devices at the outset.

The principle of acceptance for tests in the CV7000 specifications is predominantly that of assessment by attributes. Under this arrangement the conditions of sampling, measurement and limit values are laid down, and a stated number of defectives is allowed for each test. The CV7000 series now includes specifications for more than a hundred types, several of which have been in production for over two years. The evidence from these tests is a valuable addition to that obtained by manufacturers' normal life tests, particularly with regard to the mechanical and the environmental tests which are not usually carried out so often. In addition, the large numbers of devices that have to be handled give information about the incidence of inoperative failures, which in some devices occur at a rate sufficiently low as to require a substantial amount of evidence being available for its estimation.

The evidence on two types of transistor supplied to CV7000 specifications is shown in Table 1, where it will be seen that a high order of reliability is being achieved in all tests. A striking feature of these results is that in nearly 4000 specimens of the CV7001 audio-frequency germanium transistor submitted to the various tests, not a single inoperative failure occurred, and in the CV7042 high-frequency transistor only one inoperative failure occurred in nearly 4000 specimens tested. Although the results from different types of test cannot be added in this way to indicate the inoperative failure rate under, say, average conditions of electrical operation, the results give some confidence to the belief that the devices concerned are of an intrinsically reliable kind. The following Section therefore deals with the problem of assessing the inoperative failure rate for these and similar devices when used under typical conditions of operation.

Table 1

Summarized Results of Group E and Group F Tests in CV7001 and CV7003 Germanium Transistors

TEST	CV7001		CV7003	
	Tested	Failed	Tested	Failed
Temperature and Climatic Cycling	170	0	226	2
Vibration-Fatigue	170	0	226	1
Lead Fragility	162	0	226	0
Soldering	170	1	224	0
1000-hours Life	472	2	550	8†
Store 1 (−50° C)	1375	2	1570	1
Store 2 (+75° C)	1375	9	1570	4

† One inoperative failure

4. An Overall Estimate of the Inoperative Failure Rate for some Typical Transistors

It has already become customary to express failure rates as a percentage per 1000 hours and a desirable goal from the electronic computer usage point of view would be an inoperative failure rate of 0·01% per 1000 hours or lower. In practice the sample size necessary to establish a failure rate as low as 0·01% within a narrow confidence band, is prohibitive (a sample of about 50 000 would be necessary). For this reason indications of the order of magnitude of the inoperative failure rate must therefore nearly always come from field reports. Usually no exact information is available on the nature of failures in the field. Such information may contain "failures" due to genuine short or open-circuit devices, devices having been accidentally damaged, or device characteristics having drifted outside the circuit tolerance limits. In addition to the imprecise nature of field information we are faced with the fact that although laboratory life tests can provide a substantial amount of information up to the first 1000 hours, most tests are discontinued at this time and there is therefore little supporting evidence over later periods.

The whole difficulty of estimating the reliability of germanium transistors from field and from life test evidence at the present time can be briefly summarized as follows:

(1) There are insufficient data available on any one type of device. Thus information on several types will have to be pooled for laboratory data and field reports.

(2) We shall investigate whether the number of failures observed in the field is consistent with that found in the laboratory, with the objective of compounding all the information if it is permissible to do so.

(3) The failure rate to be determined will be over the first 1000 hours. Indications of whether this is true for subsequent periods of 1000 hours will be provided by reports from the field.

Table 2

Inoperative Failures in Laboratory Life Tests of Germanium Transistors to 1000 Hours

	Number Tested	Number of inoperative failures
A.F. Germanium Transistors (GET103 Series)	5275	3
H.F. Germanium Transistors (GET872 Series)	1364	2
A.F. "Reliable" Ge Transistors (CV7000 Series)	472	0
H.F. "Reliable" Ge Transistors (CV7000 Series)	550	1
Totals	7661	6

In Table 2 the findings of laboratory life tests to 1000 hours are summarized. These results are for the two types of germanium transistors considered in previous sections of the paper and include both commercial and CV versions. Six inoperative failures occurred in 7661 transistors tested of all types.

Table 3 summarizes the occurrence of inoperative failures during the first 1000 hours in 5 equipments employing a total of 21 390 transistors of 3 basically different types. The well-known χ^2 significance test shows that there is no significant difference between the number of failures observed in the laboratory and that observed in the field, because too few failures have been observed in both categories. We may therefore, in the situation as it stands, pool the data in order to estimate the failure rate. This gives a

Table 3

Inoperative Failures to 1000 Hours in Field Reports

Transistor type	Number of transistors	Type of equipment	Number of failures	Basic type category
GET114	12 000	Railway signalling	4	Audio-frequency GET103
GET4	2768	Data link	0	Audio-frequency GET103
GET104	2768	Computer	0	Audio-frequency GET103
GET872	1500	Computer	2	High-frequency GET872
GET573	2354	Lighting inverter	1	Power GET572
Totals	21 390	—	7	—

total of 29 051 transistors in which 13 inoperative failures occurred in the first 1000 hours. On this basis the best available estimate of the inoperative failure rate is 0·044% with 95% confidence limits of 0·023% to 0·073%. This result shows an inoperative failure rate at least of the same order of magnitude as the minimum requirement for computer-type applications. The evidence is drawn from a number of applications which contain differences in operating conditions, and although there is insufficient evidence to show that operating conditions have a marked effect on the incidence of inoperative failures, it is possible that there may be some reduction under low stress conditions. In the railway signalling application for example the failure rate in the first 1000 hours was 4 in 12 000 transistors and it is interesting to examine the number of faults occurring in subsequent 1000-hour periods in this equipment. These are indicated in Table 4 which also shows additional evidence beyond 1000 hours for some of the other equipments. Again there is insufficient evidence to make any definite statement, but the evidence in the first 1000 hours seems to be of the right order of magnitude for subsequent periods of 1000 hours. The evidence on the GET114 over the period 5000–9000 hours inclusive indicates an average rate of approximately 0·008% per 1000 hours. In many industrial applications of transistors an inoperative failure rate of 0·01% per 1000 hours appears to be acceptable economically.

Table 4

Inoperative Failures up to 9000 Hours in Field Reports

Transistor type	Number of transistors	Number of failures per period of 1000 hours "on" time								
		1st	2nd	3rd	4th	5th	6th	7th	8th	9th
GET114	12 000	4	2	1	3	2		3		
GET4	2768	0	2	2	2	1	Limit of present evidence			
GET104	2768	0	0	Limit of present evidence						
GET872	1500	2	1	Limit of present evidence						
GET573	2354	1	6	1	Limit of present evidence					
Totals	21 390	7	11	–	–	–	–	–	–	–

The foregoing evidence suggests that even lower failure rates may be achieved, but the main problem in this direction is that of acquiring sufficient information for assessing failure rates of as low an order. Where a claim for great reliability is one of the principal features of any new type of device, the estimation of the validity of this claim will be a considerable undertaking. This problem can be expected in many of the new types of components which are now in their infancy.

5. Acknowledgments

The authors would like to acknowledge the permission given by the Semiconductor Division of The General Electric Company Limited to publish life test information obtained on its behalf. Thanks are due also to other branches of the G.E.C. and to the Westinghouse Brake and Signal Company Limited for permission to use the field information.

Manuscript received by the Institution on 21st March 1961 (Paper No. 640).

INSTITUTION NOTICES

Birthday Honours

The Council of the Institution has congratulated Rear Admiral Kenneth Robertson Buckley (Member) on his appointment as a Knight Commander of the Military Division of the Most Excellent Order of the British Empire. Admiral Buckley has been Director of the Naval Electrical Department since 1958; his previous commands included the Naval Training Establishments of H.M.S. *Ariel* and H.M.S. *Collingwood.*

Royal Garden Party

The General Secretary of the Institution, Mr. Graham D. Clifford, and Mrs. Clifford had the honour of an invitation to the Royal Garden Party at Buckingham Palace on Thursday, 11th May, last.

Institution Meetings during Next Session

The Programme and Papers Committee is planning two symposia to be held during the 1961–62 Session. These are:

Symposium on Data Transmission—a one-day Symposium to be held in December 1961, consisting of three sessions:

Digital Data Transmission
Analogue Data Transmission
Data Collection and Distribution

The closing date for the receipt of complete manuscripts is 30th September 1961.

Symposium on Recent Developments in Industrial Electronics—a two-day Symposium to be held in April 1962, consisting of four sessions:

Measurement Techniques
Machine Tool Techniques
Process Control Techniques
Miscellaneous Electronic Techniques

The closing date for the receipt of complete manuscripts is 1st January 1962.

The Committee invites members whose professional interests lie in these fields to submit proposals for papers. Offers should include in the first instance a synopsis.

Radio Trades Examination Board

The British Radio Equipment Manufacturers' Association has nominated Mr. M. Exwood (Member) as one of its representatives on the Council of Management of the Radio Trades Examination Board.

Mr. Exwood's appointment follows the retirement of Mr. E. M. Lee (Member) owing to ill-health. Mr. Lee was Chairman of the Radio Trades Examination Board during the period 1956–1958.

Extraordinary General Meeting

A notice of a General Meeting for the election of President and Officers included advice of the Extraordinary General Meeting held on 7th June last. The recommendation and resolutions proposed by the General Council of the Institution were approved. Minutes of the meeting will be published in the July *Journal.*

Education and Training Meeting

The Education and Training Committee is organizing a meeting of heads of departments of technical colleges and training officers in industry to be held on Wednesday, 27th September at University College, London.

In the morning the meeting will receive details of the Institution's new Graduateship Examination syllabus and there will be discussion concerning the introduction of these syllabuses and the changes which will be necessary in the exempting qualifications.

After lunch there will be a symposium on practical training.

Attendance will be by invitation only; any member of the Institution who would like to receive an invitation should write to the Education Officer as soon as possible.

Recent Appointments in the Communications Field

Members in the Far East in particular will be pleased to learn that Mr. William Stubbs, C.B.E., M.C. (Member), has been appointed Secretary General of the Commonwealth Telecommunications Board. Mr. Stubbs was formerly Director General of the Federation of Malaya and the State of Singapore, a post he had held since 1957. He was elected a Member of the Institution in 1949 and he has been a representative of the Council in Malaya for a number of years.

Chief Inspector William Norman Bruce, B.E.M. (Member), has been in charge of police telecommunications with the Edinburgh City Police since 1933 and has recently taken up an appointment as Communications Officer with the Scottish Home Department. Chief Inspector Bruce has been an active member of the Scottish Section Committee for a number of years; he was elected an Associate Member in 1944 and was transferred to full Member 1949.

Completion of Volume 21

This issue completes Volume 21 of the *Journal*; Volume 22 will comprise the July to December issues. This division of the year's issues into the two volumes was referred to in the January *Journal.* The index for Volume 21 will be circulated to all members and subscribers with the August *Journal.*

U.D.C. 621.376.223

Constant-Resistance Modulators

By

Professor D. G. TUCKER,
D.Sc. (*Member*)†

Summary: It is shown how Zobel's constant-resistance networks may be interpreted in terms of constancy with time when the elements of the network are time-variable, but that the original conception of constancy with frequency when the elements are frequency-variable may be simultaneously realized in the same network. The application to rectifier modulators is examined, and it is shown that for the whole class of networks, the simplest condition for the realization of constant input resistance coincides with that for the elimination of even-order modulation products provided the carrier waveform contains only odd-order harmonics. In symmetrical networks, it also coincides with the condition for minimum conversion loss. It is shown how a constant-resistance network can be made to give a frequency-response (or output spectrum) corresponding to a specified frequency-dependent series or shunt modulator.

1. Introduction

The idea of constant-resistance modulators was introduced briefly in a previous paper,[1] but will now be studied rather more fully. Zobel[2] showed—so long ago that his name is now often not associated with the subject—that there was a class of networks which, while containing reactances and having an insertion loss varying with frequency, nevertheless had a constant-resistance iterative impedance. The principle involved was a reciprocal or dual relationship between series and shunt impedances. The main attraction of the constant-resistance iterative impedance was that several networks could be joined in cascade and their insertion losses merely added together. The same principle of reciprocal relationship can be applied[3] to modulator circuits if assumed linear—or indeed to any linear circuit with time-varying parameters[4]—and then the constant-resistance iterative impedance is constant with respect to time instead of, or as well as, with respect to frequency. There is not, of course, any immediate attraction in the idea of cascading modulator circuits; but it will be shown that it has possibilities. A more likely practical value of constant input impedance is the possibility of paralleling modulator inputs with, theoretically, no decoupling impedances.

Provided the modulator circuit is non-reactive, a constant input resistance can in some cases be obtained without it being iterative when the series and shunt elements comprise, not proper duals, but duals with a constant quantity added or subtracted. Since real rectifiers may give approximately this condition, the matter is of importance.

It will be shown, too, that when a constant input resistance is obtained in the ways mentioned above, then even-order modulation products are absent from the output of the modulator. This is of obvious practical importance, but it was shown previously[1] that even-order products can be eliminated also from the simple series or shunt modulator.

The relationship of the condition for minimum conversion loss with that for constant input resistance will also be investigated.

†Electrical Engineering Department, University of Birmingham.

2. Forms of Constant-resistance Modulator Circuit and Practical Arrangements

Figure 1 sets out the principal forms of constant-resistance network adapted for use as rectifier modulators. The time-varying resistances $r_+(t)$ and $r_-(t)$ are the linear reciprocal elements, such that

$$r_+(t) . r_-(t) = R_R^2 \qquad \ldots\ldots(1)$$

for all values of t. It can easily be checked that all these arrangements give a constant-resistance iterative impedance, but it will be found that the unsymmetrical ladder networks—(b) (i) and (b) (ii)—are iterative from left to right only, while all the others are iterative both ways.

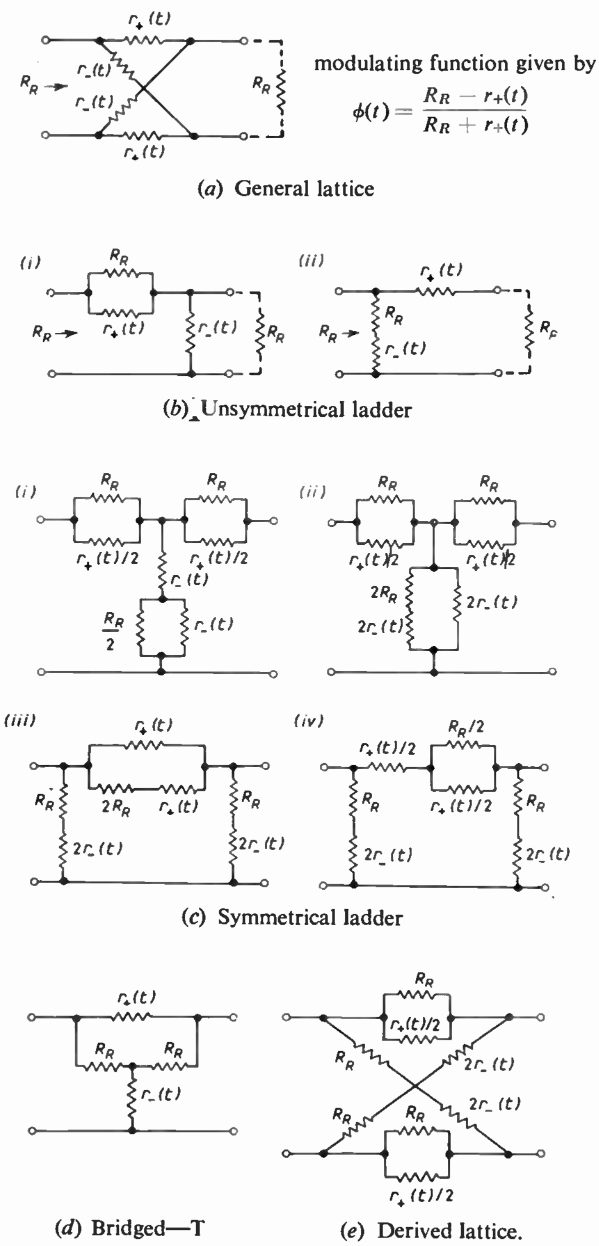
The modulating function—discussed in the previous paper—is the reciprocal of the insertion loss of the network, and is, of course, a function of time. The general lattice network, Fig. 1 (a), is the only form of network shown which can give a polarity reversal; its modulating function is

$$\phi(t) = \frac{R_R - r_+(t)}{R_R + r_+(t)} \qquad \ldots\ldots(2)$$

and clearly, for a periodic variation of $r_+(t)$, is a wave symmetrically disposed about its mean value of zero. All the other forms of network are incapable of giving a reversal of polarity; their modulating function is

$$\phi(t)=\frac{R_R}{R_R+r_+(t)} \qquad \ldots\ldots(3)$$

and this therefore varies from a large positive value to a small positive value.



For (b) to (e), modulating function given by

$$\phi(t)=\frac{R_R}{R_R+r_+(t)}$$

Fig. 1. Constant-resistance modulator forms.
(N.B. $r_+(t).r_-(t)=R_R^2$ in all cases)

The practical modulator circuit arising from the general lattice network is the well-known ring modulator, adequately described in the literature, and particularly discussed in relation to present interests in the previous paper.[1] None of the other forms converts quite so readily to a practical modulator circuit, and there is a variety of arrangements possible for each form. We shall, therefore, consider practical arrangements only for the two unsymmetrical forms, (b) (i) and (b) (ii). Two possibilities for each are shown in Fig. 2. In each case, arrangement (i) uses four rectifiers for each time-varying resistance, preserving carrier cancellation (a usual though not essential requirement) but not providing a balanced signal path. The arrangements (ii) use fewer rectifiers but more transformers, and provide a balanced signal path. Since the reciprocal time-varying resistances have to be interpreted as rectifiers driven by carrier voltages of opposite polarity (relative to the rectifier's conducting direction), we can first of all replace $r_+(t)$ and $r_-(t)$ by $r(+V_c)$ and $r(-V_c)$, and then observe that the requirement

$$r(+V_c).r(-V_c)=R_R^2 \qquad \ldots\ldots(4)$$

which must hold for all relevant values of V_c can only be met if the carrier voltage is the same (apart from polarity) across all rectifiers. This explains the necessity for the central transformer in arrangement (b) (ii) of Fig. 2. Practical arrangements for the other forms in Fig. 1 can be devised along the same lines.

It should be observed that if $r(t)$, or $r(V_c)$, contains some constant part r_d (positive or negative) which prevents the reciprocal relationship of eqn. (4) being obtained, then in network forms (a) and (b) (ii) of Fig. 1 this constant part may be removed into the terminations. In the lattice, r_d is removed from each arm and added to each termination, so that the constant-resistance condition is

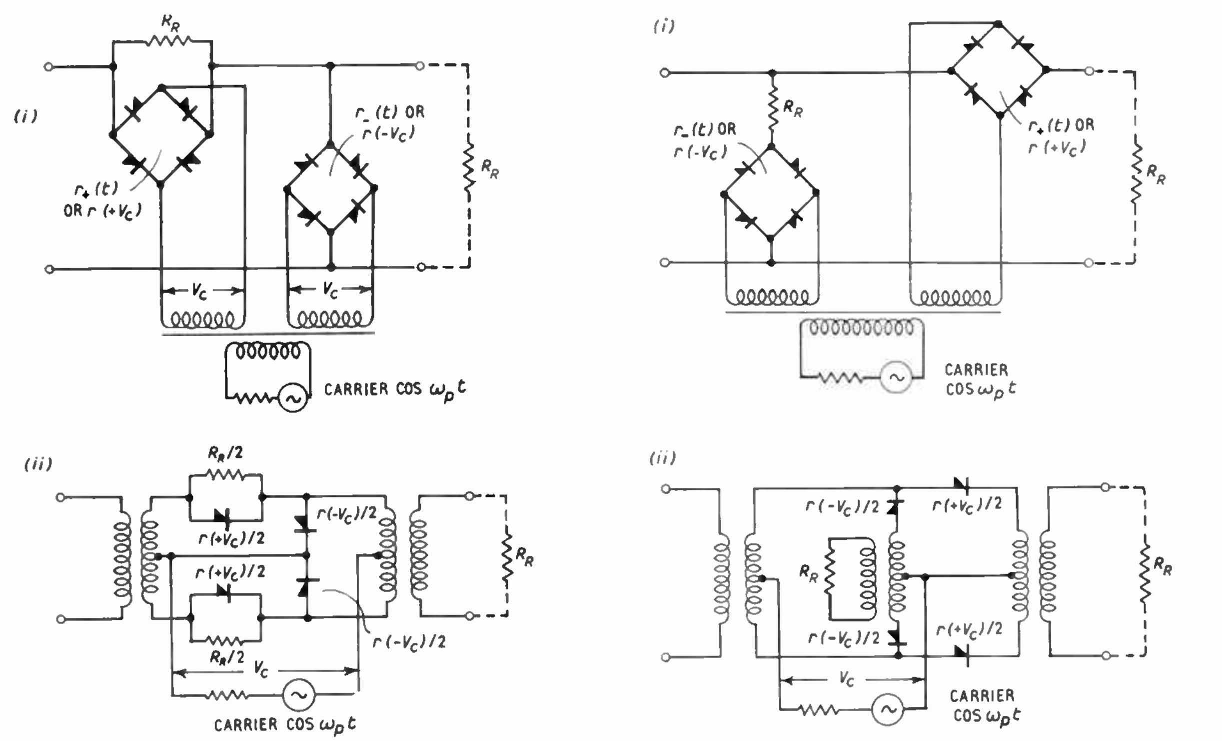
$$[r(+V_c)\pm r_d][r(-V_c)\pm r_d]=(R_R\mp r_d)^2 \qquad \ldots\ldots(5)$$

and the input resistance is $R_R\mp 2r_d$. In the ladder network (b) (ii), r_d is removed from each rectifier element directly into the R_R in series with it, so that a constant input resistance is obtained for the condition according to eqn. (5), but the input resistance is $R_R\mp r_d$.

If $g(t)=1/r(t)$, and can be similarly regarded as having a constant part g_d, then in network form (b) (i) of Fig. 1, this constant part can be removed and added to the conductance of the resistances R_R. The constant-resistance condition is then, in terms of $g(V_c)$ rather than $g(t)$,

$$[g(+V_c)\pm g_d][g(-V_c)\pm g_d]=(G_R\mp g_d)^2 \qquad \ldots\ldots(6)$$

and the input conductance is $G_R\mp g_d$, where $G_R=1/R_R$. The same process can, if desired, be



(*a*) corresponding to (*b*) (i) of Fig. 1.

(*b*) corresponding to (*b*) (ii) of Fig. 1.

Fig. 2. Practical constant-resistance modulator circuits.

applied to the lattice network, the conductance g_d being added to both the terminating conductances, and the input conductance becoming $G_R+2\,g_d$.

It should be noted that practical rectifiers often have resistance/voltage characteristics[3] which approximate to the form

$$r = r_d + a_1 \exp(-b_1 V_c) \qquad \ldots\ldots(7)$$

or conductance/voltage characteristics which approximate to the form

$$g = g_d + a_2 \exp(b_2 V_c) \qquad \ldots\ldots(8)$$

Clearly these characteristics are suitable for making constant-resistance modulators according to eqns. (5) or (6).

3. Elimination of Even-order Modulation

For the ring modulator—i.e. the lattice form (a) of Fig. 1—this matter has been adequately discussed in the previous paper.[1]

For the other modulators and *not* assuming that the reciprocal condition of eqns. (5) and (6) apply, the modulating function is of similar form to that of the series and shunt modulators discussed in the previous paper, and thus the same criterion can be used for absence of even-order modulation products, namely that the modulating function should contain no even-order harmonics. This was shown to be the case, for waveforms of V_c containing only odd harmonics of ω_p, when

$$\phi(+V_c)+\phi(-V_c) = K \qquad \ldots\ldots(9)$$

for all values of V_c in the range concerned; K is a constant independent of V_c, and the modulating function is expressed as a function of V_c rather than of time. Now when we do not assume the iterative condition, the calculation of $\phi(V_c)$ usually becomes rather complicated, and to avoid tiresome repetition here, we shall consider only one case—(b) (ii) of Fig. 1—in detail, and merely state the results for the others.

For the circuit (b) (ii) of Fig. 1, and its practical arrangements of Fig. 2 (b), it may be shown that

$$\phi(+V_c) = \frac{R_S+R_R}{R_R+r(+V_c)} \times \frac{[R_R+r(+V_c)][R_R+r(-V_c)]}{R_R^2+2R_SR_R+(R_S+R_R)[r(+V_c)+r(-V_c)]+r(+V_c).r(-V_c)} \qquad \ldots\ldots(10)$$

so that

$$\phi(+V_c)+\phi(-V_c)=\frac{(R_S+R_R)[2R_R+r(+V_c)+r(-V_c)]}{R_R^2+2R_SR_R+(R_S+R_R)[r(+V_c)+r(-V_c)]+r(+V_c).r(-V_c)} \quad ......(11)$$

$= K$ from eqn. (9).

If we put $K = 1$, we get very quickly the result

$$r(+V_c).r(-V_c) = R_R^2 \quad ......(12)$$

which must hold for all relevant values of V_c; and suitable manipulation enables us to see that if we put

$$(R_S+R_R)\left(1-\frac{1}{K}\right)=\pm r_d \quad ......(13)$$

then we obtain the condition

$$[r(+V_c)\pm r_d][r(-V_c)\pm r_d] = (R_R\mp r_d)^2 \quad ...(14)$$

where r_d is regarded as a positive resistance.

This condition is identical with eqn. (5), i.e. with the condition for constant input resistance.

It will be seen that the signal-source resistance, R_S, drops out of the condition altogether, except that as r_d must be a pure resistance, then, from eqn. (13), R_S must be a pure resistance; its value is immaterial since K may be chosen arbitrarily.

The principle of duality makes it clear that for network (b) (i) of Fig. 1, the condition corresponding to eqn. (14) for the elimination of even-order modulation is that given already for constant input resistance in eqn. (6).

That the simpler condition (12) applies to all forms of the network in Fig. 1 (apart from (a)) is clear from the fact that this condition makes the networks iterative, and when iterative, they all have the same modulating function. The more complex conditions, e.g. eqn. (14), do not apply to forms other than those for which they are worked out.

4. Minimum Conversion Loss

For a modulator with resistive terminations, the conversion loss is inversely proportional to the fundamental component of the modulating function (assuming it is first-order modulation which is being considered). Therefore, in all normal practical circumstances, and with a symmetrical waveform of V_c, i.e. one containing only odd harmonics, the conversion loss is a minimum when $\phi(+V_c)-\phi(-V_c)$ is a maximum for all relevant values of V_c. Since we wish to maximize this function with respect to R_S and R_R (and not V_c), we shall designate it merely as ϕ.

The lattice, i.e. the ring modulator, was dealt with in reference 1 and led to the result that ϕ was a maximum when *either* R_S *or* R_R was equal to $\sqrt{[r(+V_c).r(-V_c)]}$ for all values of V_c. This result arises from the fact that a symmetrical 2-terminal-pair network of any form gives its minimum insertion loss when it is iteratively terminated in at least one direction. Since the network is symmetrical this is the same as saying it is image-terminated at least at one end. A general proof of this is not difficult, but will not be given here, as some discussion of it is available in text-books.[5] The consequence of this property of symmetrical networks is that all the constant-resistance modulators shown in (c), (d) and (e) in Fig. 1 have a minimum conversion loss when $r(+V_c).r(-V_c) = R_R^2$, and so can have the three conditions (a) constant input resistance, (b) no even-order modulation and (c) minimum conversion loss, all occurring together.

The unsymmetrical modulators of Fig. 1 (b), however, are different. Consider the modulator (b) (ii), and assume it has to work between matched source and load resistances, i.e. $R_S = R_R = R$ (say). Then the function ϕ is easily shown to be

$$\phi=\frac{2R[r(-V_c)-r(+V_c)]}{3R^2+2R[r(+V_c)+r(-V_c)]+r(+V_c).r(-V_c)} \quad ......(15)$$

This has a maximum value when $d\phi/dR = 0$, i.e. when

$$R^2 = r(+V_c).r(-V_c)/3 \quad ......(16)$$

so that the minimum conversion loss occurs when the terminations are $1/\sqrt{3}$ of the iterative impedance.

The modulator of (b) (i) in Fig. 1 can similarly be shown to give its minimum conversion loss when

$$R^2 = 3r(+V_c).r(-V_c) \quad ......(17)$$

The actual calculation of conversion loss for any of these modulators is rather complicated except when $r(t)$ is a square wave. In general, recourse has to be taken to the methods of infinite equations discussed in other papers[6,7,8]; but as these require the Fourier expansion of $r(t)$ in harmonics of ω_p, it may sometimes, in the purely resistive circuits at present under discussion, be just as easy to find the Fourier expansion of $\phi(t)$ and so obtain the conversion loss directly. The matter will not be discussed further here.

5. Constant-resistance Modulators with Frequency-dependent Circuits

Since the circuits we are considering are all linear, there is no reason why reactances should not be included even though the resistance is time-varying. (Indeed, Zadeh[4] shows that time-varying reactances may be incorporated.) An example is given in Fig. 3,

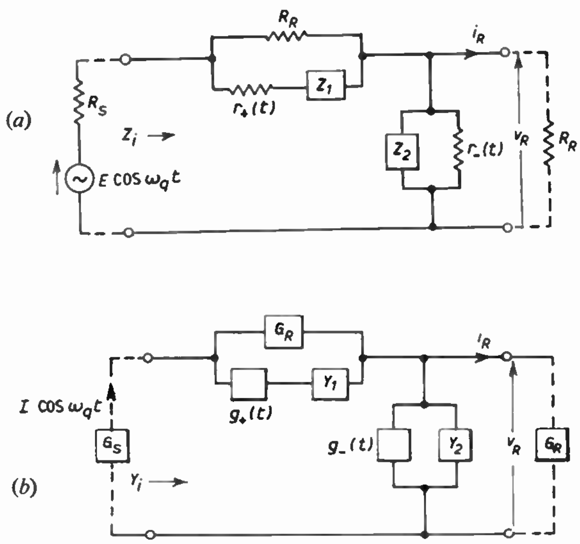


Fig. 3. Constant-resistance modulator with reactances.

(a) $Z_i = R_R$ if $r_+(t).r_-(t) = R_R^2$ and $Z_1 Z_2 = R_R^2$

(b) $Y_i = G_R$ if $g_+(t).g_-(t) = G_R^2$ and $Y_1 Y_2 = G_R^2$

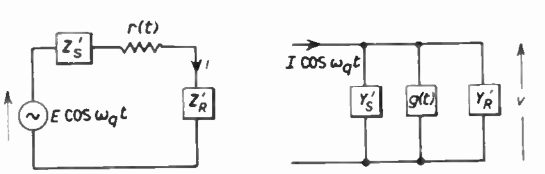
where the network of type (b) (i) of Fig. 1 is shown with reactances, and drawn in terms of impedances in (a) and admittances in (b). This network, under the conditions stated on the figure, gives an iterative impedance which is a pure resistance, constant with both time and frequency. The modulation, however, is frequency-dependent. Since the modulating function is as given by eqn. (3) when the network is iterative, then the output current, i_R, is evidently given in terms of Fig. 3 (a) by

$$i_R = \frac{V \cos \omega_q t}{R_R + Z_1 + r_+(t)} \qquad \text{......(18)}$$

where V is a constant. Similarly the output voltage, v_R, may be written, from Fig. 3 (b),

$$v_R = \frac{I' \cos \omega_q t}{G_R + Y_2 + g_-(t)} \qquad \text{......(19)}$$

where I' is a constant.



(a) Series modulator (b) Shunt modulator

Fig. 4. Series and shunt modulators with same output spectrum as Fig. 3 (a) and Fig. 3 (b) respectively.
(Relations discussed in text.)

Now Fig. 4 (a) shows a series modulator in terms of impedances, and Fig. 4 (b) shows a shunt modulator in terms of admittances. In the series modulator, the current flowing through the load is

$$i = \frac{E \cos \omega_q t}{Z'_S + Z'_R + r(t)} \qquad \text{......(20)}$$

Comparing this with eqn. (18), we see that if

$$R_R + Z_1 = Z'_S + Z'_R \qquad \text{......(21)}$$

then the constant-resistance modulator has exactly the same frequency response of output current as the series modulator, but in addition has a constant input resistance, which the series modulator does not have.

In the shunt modulator, the voltage across the load is

$$v = \frac{I \cos \omega_q t}{Y'_S + Y'_R + g(t)} \qquad \text{......(22)}$$

Comparing this with eqn. (19), we see that if

$$G_R + Y_2 = Y'_S + Y'_R \qquad \text{......(23)}$$

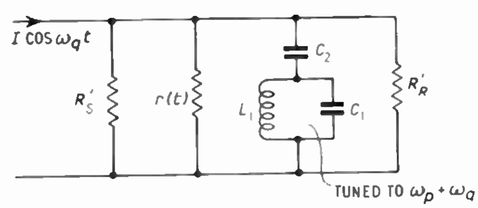


Fig. 5. Shunt modulator tuned to eliminate unwanted sidebands.

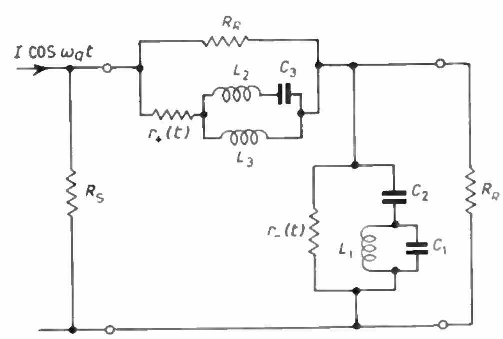


Fig. 6. Constant-resistance modulator giving same effect as modulator of Fig. 5.
$r_+(t).r_-(t) = R_R^2$; $L_2 = C_1 R_R^2$; $C_3 = L_1/R_R^2$; $L_3 = C_2 R_R^2$

then the constant-resistance modulator has exactly the same frequency response of output voltage as the shunt modulator.

A practical example of the application of these principles is shown in Figs. 5 and 6. Figure 5 shows a shunt modulator with a tuned output which, while permitting the voltage to develop at the wanted sideband frequency ($\omega_p+\omega_q$, say), provides a short circuit (effectively) at all other sideband frequencies. This sort of arrangement would often be used in practice. The capacitance C_2 is necessary to avoid a short circuit at the signal frequency (ω_q), which is here assumed to be much lower than ω_p. It would, however, not be possible to parallel a number of different modulators without interference, since their input impedances are both time- and frequency-dependent. If the equivalent constant-resistance modulator is made, however, as shown in Fig. 6, the same tuning of the wanted sideband is obtained, but a number of modulators may now be paralleled since the input impedance is just a constant resistance.

It is possible that where double and multiple modulation is needed, advantage would accrue from the use of two or more constant-resistance modulators in cascade, rather than two or more series or shunt modulators separated by a decoupling attenuator or amplifier. In many types of electronic system the power supply and maintenance of the amplifier represent difficulties and the alternative decoupling of an attenuator would greatly worsen the noise factor in a low-level system. Cascaded constant-resistance modulators with sideband selection built in to them as in Fig. 6 would be worth while in these cases.

6. References

1. D. G. Tucker, "Elimination of even-order modulation in rectifier modulators", *J.Brit.I.R.E.*, **21**, p. 161, February 1961.
2. O. J. Zobel, "Distortion correction in electrical circuits with constant resistance recurrent networks", *Bell Syst. Tech. J.*, **7**, p. 438, 1928.
3. D. G. Tucker, "Rectifier resistance laws", *Wireless Engineer*, **25**, p. 117, 1948.
4. L. A. Zadeh, "Constant-resistance networks of the linear varying-parameter type", *Proc. Inst. Radio Engrs*, **39**, p. 688, 1951.
5. E. A. Guillemin, "Communication Networks", Vol. 2, Chapter 2, Section 17 (John Wiley, New York, 1935).
6. D. P. Howson and D. G. Tucker, "Rectifier modulators with frequency selective terminations, with particular reference to the effect of even-order modulation products", *Proc. Instn Elect. Engrs*, **107B**, p. 261, May 1960. (I.E.E. Paper No. 3051E, January 1960.)
7. D. G. Tucker, "Rectifier Modulators with Time-varying Resistance Function containing Even Harmonics of Carrier Frequency", Electrical Engineering Department, University of Birmingham, Memorandum No. 47, 1959.
8. D. P. Howson, "Rectifier modulators: analysis by successive approximations", *Electronic Technology*, **37**, p. 158, April 1960.

Manuscript received by the Institution on 18th March 1960 (Paper No. 641.)

U.D.C. 621.385.832 : 681.142

A Cathode-ray Tube Output for a Digital Computer

By

R. L. GRIMSDALE, M.Sc., Ph.D.†

AND

M. BARRACLOUGH, B.Sc.†

Presented at the Computer Group's Symposium on Alpha-Numeric Display Devices held in London on 18th January 1961.

Summary: The unit gives a visual display on one cathode-ray tube and has a second tube which is equipped with an automatic camera controlled directly by the computer. The device may be used for curve plotting with a resolution of 256 × 256 or for alpha-numeric display. In the latter case the computer can be programmed to display a variety of different formats and high output rates are possible.

1. Introduction

The results of many scientific and engineering computations are required in graphical form. The use of the normal output channel of a digital computer followed by manual plotting is both slow and tedious. A direct means of plotting results from a computer is thus of great value particularly if the plotted curves can be inspected visually as the calculation proceeds, at the same time being photographed for a permanent record.

The results are displayed simultaneously on two standard 9 in. electrostatic deflection cathode-ray tubes, one of these being for direct visual display and the other used in conjunction with an automatic camera. The unit is controlled by a special computer instruction which causes a point to be displayed within a matrix of 256 × 256 possible positions. The equipment has been designed for use in conjunction with the Mercury computer and some special features have been incorporated to achieve an economical circuit arrangement compatible with the special features of that machine.

There is no restriction on the form of display and alpha-numeric characters can be displayed by a suitable programme. The unit can thus be used as a fast printer for tables of results or for writing information on graphs.

2. General Description

2.1. *Resolution and Accuracy*

The choice of a resolution of 256 × 256 points was primarily governed by the number of points which can be resolved in the working area of the cathode-ray tube screen. This resolution is moreover well within the limits to which the accuracy and stability of the associated circuits can be designed.

The stability of the unit over moderately long periods is important in the cases where a photographic record is being obtained by displaying points as they are calculated. To avoid drift, circuits employing negative feedback are used throughout. The linearity and orthogonality of vertical and horizontal lines is determined in the manufacture of the cathode-ray tube and defects of this type in available tubes have not produced observable errors. In cases where these errors may be important it is convenient to plot calibrated axes or graticules at the same time as the graph is plotted. Where very high accuracy is required alternative display devices could be employed.[1]

2.2. *The Supply of Information from the Computer*

The design of the unit is influenced by the way in which the coordinate values of the points are obtained from the Mercury computer.[2] The arithmetic unit of this machine operates in the serial mode with a 1 Mc/s clock rate, but the main store which uses a matrix of magnetic cores, works in the parallel mode. Outputs from this store are available on ten parallel channels for about 8 μs of a 10 μs period. As each coordinate has 256 possible values it requires 8 bits for its specification. The arrangement employed is to use eight out of ten bits of one word for the *X*-coordinate and a similar number of bits out of the succeeding word for the *Y*-coordinate. Thus the two coordinates are specified in two successive 10 μs periods and some form of intermediate storage is required. The use of a set of staticizers is uneconomical and the system to be described is preferred.

2.3. *The Analogue Coordinate Stores*

In the system employed the coordinate values are converted into an analogue representation and stored in this form. As the digital forms of the two coordinates occur at different times only one digital-to-analogue converter circuit is required. The overall scheme is

†Electrical Engineering Laboratories, University of Manchester.

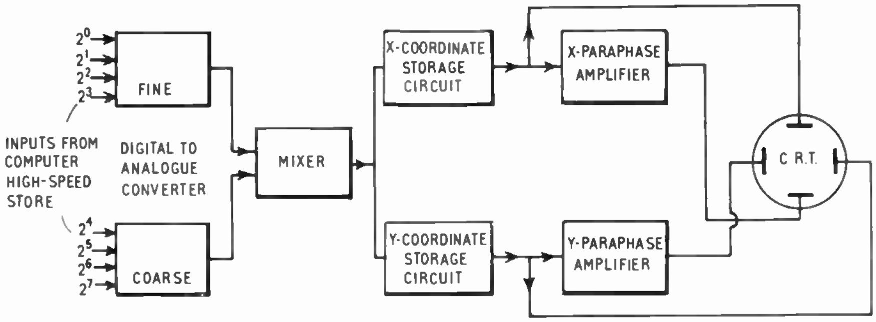


Fig. 1. Block diagram of the cathode-ray tube display unit.

shown in Fig. 1. In the first 10 μs period the *X*-coordinate is converted and stored on the *X*-coordinate store. In the next 10 μs the same process is applied to the *Y*-coordinate. The outputs of the analogue storage circuits are directly connected to the paraphase amplifiers for the cathode-ray tube, and the selected point is displayed during some subsequent period by applying a bright-up waveform of 20 μs duration to the grid of the tube.

3. Design Details

3.1. *The Digital-to-Analogue Converter*

The digital-to-analogue converter employs the conventional technique of current summation by an anode-follower. Each of the eight bits is represented by a current bleed which is controlled by a diode gate. As the ratio of the current for the most significant to that for the least significant is 128 to 1, the converter is divided into two stages (Fig. 2). The current bleeds provided by resistors R1, R2, R3, R4 have the ratios 1 : 2 : 4 : 8 and the second converter using V2 is arranged in an identical manner. The outputs from the two units are mixed in the ratio 16 to 1 by the anode follower V5. The complete converter is designed to operate in under 2 μs and to achieve this speed and to give low output impedances, cathode followers (V3, V4 and V6) are employed in the feedback circuits of the anode followers. The output channels from the

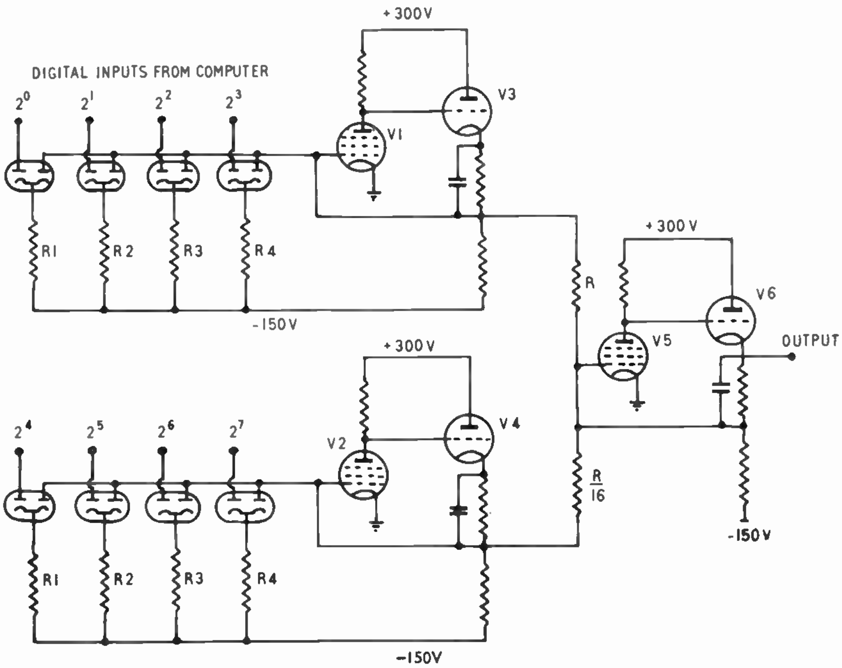


Fig. 2. The digital-to-analogue converter.

core store are permanently connected to the digital-to-analogue converter and the output from the converter is similarly connected to input circuits of both the *X* and *Y*-coordinate analogue stores.

3.2. *The Analogue Coordinate Stores*

Identical storage circuits are provided for holding the *X* and *Y*-coordinate values. The analogue quantity is stored as a charge on a capacitor in a Miller-type circuit and this circuit is included in an overall feedback network during the setting process. This arrangement ensures that the store is set to a high accuracy in the short time available. The digital input to the unit from the computer is available for 8 μs of which 2 μs are allocated for digital-to-analogue conversion. Allowing for a margin of safety, the time for correct setting of the storage circuit is 4 μs. The integrator. Assuming the gain from the output of the storage circuit V5 to the output of the cathode follower, V3, is G and v_1 and v_0 are respectively the input and output voltages then a current $\frac{G(v_1 - v_0)}{R^*}$ will flow into the grid circuit of the Miller circuit. Hence

$$\frac{G(v_1 - v_0)}{R^*} = \frac{C\,\mathrm{d}v_0}{\mathrm{d}t}$$

where R^* is the effective input resistance (formed by R3+R4 or by R5+R6) and C is the capacitance of the Miller capacitor. The error therefore reduces exponentially with a time-constant R^*C/G which is approximately 0·05 μs. The time in the non-linear condition will not exceed 1 μs in the worst case; hence in the remaining 3 μs the final error in setting is negligible compared with ½ V which is the voltage change corresponding to one out of the 256 steps.

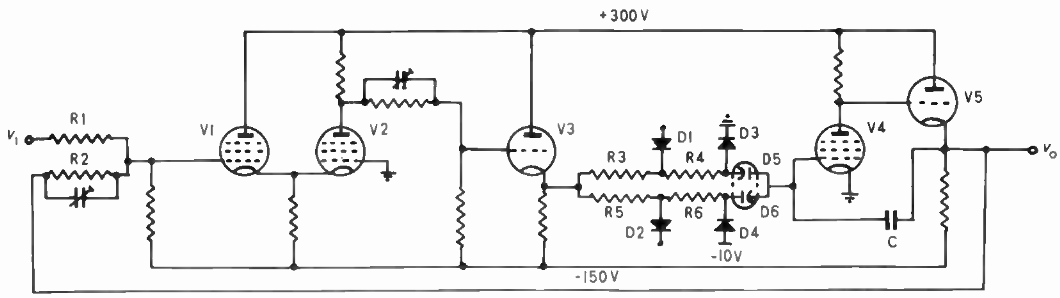


Fig. 3. The analogue coordinate store.

circuit is shown in Fig. 3; V4 and V5 form the Miller storage circuit which has a gated input under the control of the six diodes which form a bi-directional switch. During a setting period the switch is closed and the input voltage is applied via the long-tailed pair (V1 and V2) and the cathode follower (V3) to the storage circuit. There is a feedback path from the output of the storage circuit to the input of the unit via the resistance R2.

The circuit is designed to operate non-linearly when there is more than a small difference (about 6 V) between the input value and the output or stored value. Under these circumstances V1 or V2 is cut-off and a constant current flows into or out of the Miller storage circuit through one of the two thermionic diodes. This arrangement permits rapid charging or discharging to take place without reducing the available gain with linear operation.

In the linear regime the feedback loop is completed and the circuit may be considered in two parts. One part comprises the long-tailed pair and cathode follower; this circuit gives gain without inversion and with the introduction of only one important lag. The second part, the Miller storage circuit, forms an

At the end of the setting period the bi-directional switch is opened by turning on diodes D1 and D2 which in turn cut off diodes D5 and D6. The cathode of D5 is then held slightly above earth and the anode of D6 is held slightly below −10 V. The purpose of diodes D3 and D4 is to establish the potentials of the relevant electrodes of D5 and D6 at these defined potentials rather than to allow the possibility of fluctuations in the levels of the controlling waveforms on diodes D1 and D2 being transmitted to the Miller circuit via the interelectrode capacitance of diodes D5 and D6.

During the storage period there is negligible loss of charge through the cut-off diodes D5 and D6 and the leakage due to grid current in the pentode is reduced to a negligible amount by the use of a 6BR7 type operating at low anode current. Measurements obtained by extending the period of storage show that the loss of charge during the normal storage period is insignificant.

The outputs of the *X* and *Y*-coordinate storage circuit are connected directly to the respective paraphase amplifiers driving the cathode-ray tube plates.

When both the storage circuits have been set the required point is displayed by applying a bright-up pulse to the grid of the cathode-ray tube for a period of 20 μs.

4. Applications

The unit can either be viewed directly or the results may be photographed. In the latter case the shutter of the camera is held open for the duration of the complete picture and each point is displayed once. The camera is under the control of the computer, one instruction opening the shutter and another closing the shutter and moving the film on for the next exposure.

There is no limit to the number of points which may be displayed on one exposure of the film nor is there a limit to the total duration of the display so that the points may be displayed as they are calculated. Standard 35 mm oscillograph recording film is used.

An example of a method of displaying numerals is illustrated diagrammatically in Fig. 4. In this the dots represent available points while the crosses represent points actually displayed. The characters each occupy an 8 × 8 square of points and are seven points high and up to six points wide. An average of sixteen points are required for each character. The time to display each character is thus 2·88 ms but further time is required to add constants to the coordinates to produce the required X and Y shifts. It is possible to produce discernible numerals with an average of only ten points per character. Figure 5 shows a page of results using numerals of the type shown in Fig. 4. The total time to produce this, including binary to decimal conversion is 3·5 seconds. In contrast with the teleprinter normally used for printing results no time is wasted in producing the spaces for a desirable layout. If larger displays and well-formed alphabetic characters are required an overall character size of 11 × 10 or 15 × 14 may be used.

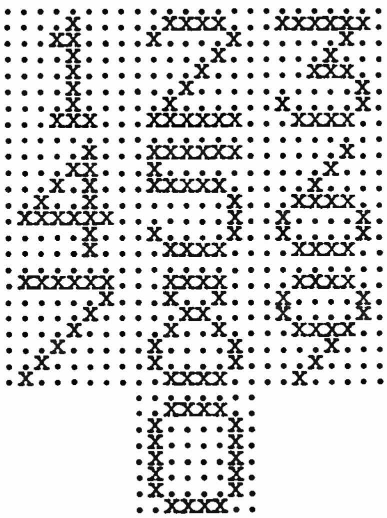
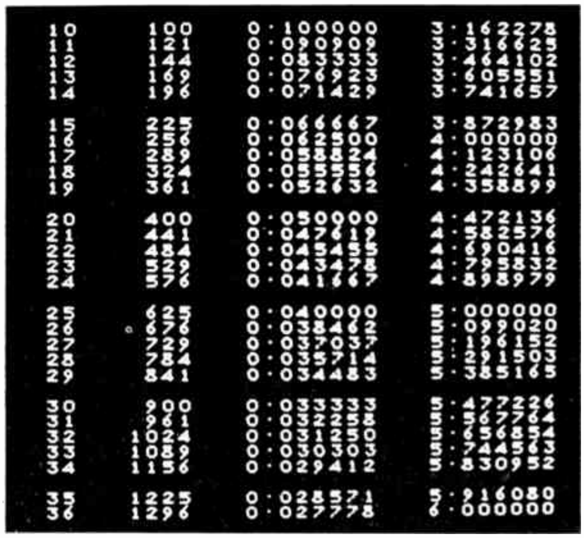


Fig. 4. Method of displaying numerals.



10	100	0·100000	3·162278
11	121	0·090909	3·316625
12	144	0·083333	3·464102
13	169	0·076923	3·605551
14	196	0·071429	3·741657
15	225	0·066667	3·872983
16	256	0·062500	4·000000
17	289	0·058824	4·123106
18	324	0·055556	4·242641
19	361	0·052632	4·358899
20	400	0·050000	4·472136
21	441	0·047619	4·582576
22	484	0·045455	4·690416
23	529	0·043478	4·795832
24	576	0·041667	4·898979
25	625	0·040000	5·000000
26	676	0·038462	5·099020
27	729	0·037037	5·196152
28	784	0·035714	5·291503
29	841	0·034483	5·385165
30	900	0·033333	5·477226
31	961	0·032258	5·567764
32	1024	0·031250	5·656854
33	1089	0·030303	5·744563
34	1156	0·029412	5·830952
35	1225	0·028571	5·916080
36	1296	0·027778	6·000000

Fig. 5. Page of results using numerals of the type shown in Fig. 4.

When a visual display is required the complete display must be replotted continuously. In order to avoid flicker the number of points per frame must be limited to obtain a rate of about 25 frames per second at least. In special cases interlace techniques may be employed to reduce flicker. For this application all the coordinates of all the points to be displayed should already be stored in consecutive locations of the high-speed store. The points can then be displayed by a special programme consisting of only two instructions, the display instruction of 120 μs and a counting instruction of 60 μs. The time to display one point is thus 180 μs and with reasonable frame flicker rate about 250 points may be displayed per frame which is adequate for most graphical displays. Similarly about 16 alpha-numeric characters can be displayed simultaneously without flicker, although this number can be doubled by employing interlacing techniques.

Figure 6 shows an example of a graphical display produced by the unit. This consists of two cosine waves with exponentially decaying amplitudes together with an identifying number in the top left corner.

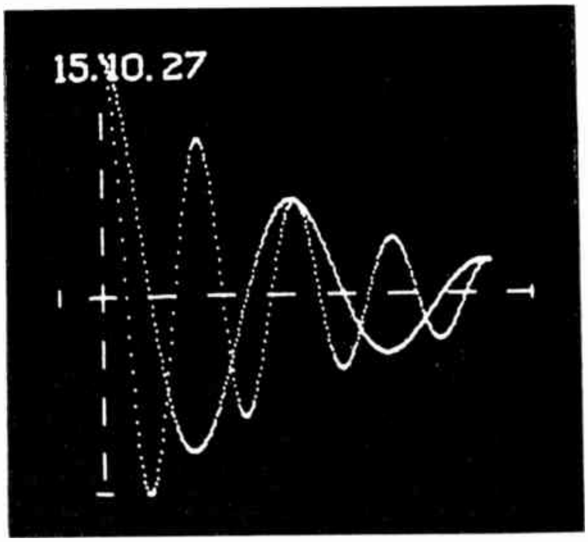


Fig. 6. Graphical display produced by the cathode-ray tube display unit.

5. Conclusions

The unit which has been in use for a period of over two years has demonstrated that a display device of this type is a valuable adjunct to a high-speed digital computer, not only for producing rapid permanent records of alpha-numeric and graphical information but as an immediate output device for visual display. As an example of the latter application it is possible to operate a digital computer in the same way as an analogue machine, permitting the results of parameter adjustments to be displayed on the screen.

6. References

1. D. C. Jeffreys and G. R. Hoffman, "The electroluminescent matrix for pattern display and recording". Paper read at the Symposium on Alpha-Numeric Display Devices. *J. Brit.I.R.E.*, **21**, 1961 (To be published).
2. K. Lonsdale and E. T. Warburton, "Mercury: a high-speed digital computer", *Proc. Instn Elect. Engrs*, **103B**, p. 175, March 1956.

Manuscript received by the Institution on 7th March 1961. (Paper No. 642.)

Progress in the Scout Satellite Programme

It was agreed some time ago that the United Kingdom would take advantage of the offer made by the National Aeronautics and Space Administration of the United States of America to launch British designed and made instruments in a satellite using Scout rockets. A total of three launchings was agreed upon and the British National Committee on Space Research, set up by the Royal Society, decided on the various physical quantities of interest in the upper atmosphere which should be measured. The first satellite, U.K.1, will be launched in late 1961 or early 1962.

The experiments for U.K.1 have been designed by three University departments of physics—Birmingham, Imperial College and University College, London, and the instruments are now ready for testing and despatch to the United States. There the instruments will be integrated with the auxiliary gear in the satellite shell which, together with the power supplies and the telemetry system has been designed in America.

Among the physical quantities which are to be measured is the concentration of electrons, using an equipment designed at the University of Birmingham under Professor J. Sayers, Dr. R. L. F. Boyd of University College, London, will measure electron concentration by a different method, and his experiment will also measure the average energy of electrons, the intensity of x-rays and ultra-violet rays and the nature of positively charged particles. This experiment will be described by Dr. Boyd in a paper at the Brit.I.R.E. Convention entitled "The Use of Probing Electrodes in the Study of the Ionosphere". The contribution of the Imperial College team, lead by Professor H. Elliot, will be in the measurement of heavy components of cosmic radiation, that is, charged carbon atoms and heavier particles. A paper on "Cosmic Ray Measurements in the U.K. Scout and Satellite", by Professor Elliot, members of his staff and an engineer with the radio manufacturer who has built the equipment, will also be given at the Convention, in Session 4.

It has now been announced that the second Scout which will be launched approximately twelve months after U.K.1 will put into orbit a satellite containing a further three sets of experiments. The first, sponsored by the Meteorological Office under Dr. K. Frith, will be concerned with the measurement of the concentration of ozone at high altitudes. This is carried out by determining the variation of ultra-violet light with the thickness of the atmosphere through which the radiation passes, thus giving a direct indication of the ozone absorption of u.v. rays.

The second experiment under Dr. F. Graham Smith of the Mullard Radio Astronomy Observatory at Cambridge is designed to extend the range of radio astronomy to study of wave-lengths below 30 Mc/s and to determine the absorption by the ionosphere of signals in the area of $\frac{3}{4}$ to 3 Mc/s at heights between 500 and 1500 km. This will in turn give an indication of the electron density and electron energy. Dr. Smith is to read a paper on "Radio Astronomy from Rockets and Satellites" at the Convention which will deal with this equipment.

The third experiment is designed by Dr. R. C. Jennison of Jodrell Bank and will study micro-meteorites, and in particular the size of the particles. Hitherto experiments on micro-meteorites have been concerned with concentration. This particular experiment involves the measurement of the size of a hole in aluminium tape punched by the particles by observing the signal produced on a cell behind the tape when the hole is illuminated by the sun. The micro-meteorites that are being studied are 0·0001 in. in diameter and have a velocity of about 100000 miles per hour.

The U.K. satellites will be launched at an angle of about 50 deg to the equator from the N.A.S.A. launching site at Wallops Island, West Virginia, and will have orbits lying between 300 and 2000 km.

For the third Scout satellite no definite research programme has yet been decided, and it may depend to some extent on the results of the first satellite.

Quite apart from the satellite programme there is continuing research work being carried on with *Skylark* and *Black Knight* rockets, and all the equipment which is being used in the satellites has been tested using these means.

Meanwhile, as British research effort on space projects increases, Industry and Government are beginning to give serious consideration to space communications. Developments in satellite communications and the questions arising were discussed at a meeting on 17th May between representatives of the Post Office and of the Society of British Aircraft Constructors, the British Constructional Steelwork Association, the Electronic Engineering Association, the British Electrical and Allied Manufacturers Association, the Radio Industry Council, and the Telecommunication Engineering and Manufacturing Association.

There was a full exchange of views on the many aspects of the subject, including the machinery that might best be adopted to foster full co-operation between the Post Office and British industry in this new and rapidly developing field.

All these activities give added point to the relevance of the 1961 Brit.I.R.E. Convention.

U.D.C. 621.396.822.029.64 : 621.317.32

Techniques of Microwave Noise Measurement

By

B. G. BOSCH, Dipl.Ing., Ph.D.†

AND

W. A. GAMBLING, B.Sc., Ph.D.‡

(*Associate Member*)

Summary: After a brief outline of some aspects of microwave oscillator noise the paper describes various techniques of noise measurement. The latter include a.m. and f.m. direct detection and superheterodyne methods as well as systems for the measurement of correlation.

1. Introduction

1.1. *The Different Noise Components in Microwave Oscillators*

A monochromatic oscillator does not exist in nature, and noise is always present in addition to the required carrier. Oscillator noise consists of two main components. The first component, normally termed background noise, is generated at microwave frequencies and corresponds to the noise encountered in linear microwave amplifiers where the noise power spectrum follows roughly the gain-frequency characteristic of the tube. The background noise is completely uncorrelated with the carrier.

The second noise component can be described as modulation noise. Due to the necessarily non-linear nature of an oscillator, this type of noise appears at microwave frequencies as a result of modulation of the signal by noise originating at lower frequencies. In this manner, frequency- as well as amplitude-modulation noise is generated. The a.m. and f.m. effects are correlated if they are produced by the same modulating event.

The output of a practical device may be represented in various ways, two of which have proved particularly useful. One uses a frequency-domain model in which the carrier is defined as the centre frequency of the output spectrum, and the other uses the time-domain model of a monochromatic carrier, amplitude and frequency modulated by noise, together with some uncorrelated background noise.[1, 2] The type of representation employed depends largely on the particular application of the oscillator.

1.2. *The Various Noise Sources in Beam-type Microwave Tubes*[3, 4]

The modulation and background noise experienced in microwave oscillators originate from a number of noise sources which are common to most types of microwave tube employing an electron beam. The more important ones will be mentioned briefly.

The main source of noise is the cathode. The random emission of electrons and the distribution of emission velocities result in current and velocity fluctuations in the electron beam, and these make a large contribution to both background and modulation noise. An additional cathode effect is flicker noise which is greatest at low frequencies and thus gives rise to modulation noise components.

Partition noise due to the random collection of electrons by the electrodes may also contribute to modulation and background noise. This effect, however, can be considerably reduced if care is taken in the design and mechanical alignment of the tube. In addition induced partition noise can occur.

Positive ions may also play a part in noise generation, particularly in tubes employing long electron beams. In microwave oscillators they seem to contribute mainly to the modulation components and manifest themselves as discrete oscillations and broad noise peaks. Noise arising from ion-electron collisions can also be present.

Lastly, an undesirable output may be produced by microphony, interference, and power supply ripple. These effects, however, can be eliminated if suitable precautions are taken, and will not be considered further.

1.3. *Effect of Oscillator Noise on Microwave Systems*

There are various ways in which oscillator noise can have an adverse effect on the performance of a microwave system.[5, 6] For example, in microwave superheterodyne receivers noise generated by the local oscillator may reduce the receiver sensitivity. The a.m. and background noise in a bandwidth equal to the intermediate-frequency (i.f.) bandwidth, and separated from the local oscillator centre frequency by the intermediate frequency, enters the i.f. amplifier together with the signal and thus increases the receiver noise figure. The problem becomes relatively more

† Telefunken GmbH, Ulm/Donau, Germany
‡ Department of Electronics, University of Southampton

serious with the availability of low-noise mixer crystals.

Noise in transmitting tubes, or tubes used as master oscillators, restricts the performance of c.w. radar systems. Here the receiver is tuned to a frequency separated from that of the transmitter by the Doppler frequency. As complete isolation between transmitter and receiver can never be achieved, the receiver can pick up noise generated by the transmitter tube in the particular band to which the receiver responds. In this case the total transmitter noise contained in this bandwidth—both modulation components plus the background component—will normally affect the receiver.

As a further example, a deterioration in performance due to transmitter noise can occur in radar equipments employing integration systems. Here the frequency deviation of the carrier, caused by f.m. noise, and the rate of deviation, influence the performance adversely.

In the first two examples quoted the frequency-domain model is of advantage in representing the noisy oscillator, whereas in the third case the time-domain model must be used since, ideally, information about the signal as a function of time has to be provided.

1.4. *Noise Quantities of Interest*

To determine the effect of noise on the application of an oscillator in a particular system, the investigation of one or more of the various noise components may be necessary. On the other hand, to obtain a better understanding of the mechanisms of noise generation a more detailed study is desirable.

The main quantities of interest are as follows:

(*a*) Total noise power spectrum.

(*b*) A.m. noise power spectrum.

(*c*) F.m. noise power spectrum.

(*d*) Background noise power spectrum.

(*e*) R.m.s. amplitude deviation.

(*f*) R.m.s. frequency deviation.

(*g*) Frequency deviation spectrum.

(*h*) Correlation between a.m. and f.m. noise.

In addition it may be necessary to separate the spectra into upper and lower sidebands. Items (*a*), (*b*), (*c*), (*d*) and (*h*) refer to the frequency-domain model and (*e*), (*f*), (*g*) and (*h*) to the time-domain model.

It should be mentioned that all the noise measuring methods to be described can equally well be applied to microwave amplifiers carrying a signal where, in addition to background noise, modulation noise can arise through operation in the non-linear region.

2. Basic Techniques

There are certain principles which form the basis of all the noise measuring methods discussed here, and these will be outlined before the individual methods are described in detail. As might be expected, and as shown by previously published work, the noise power spectra encountered in microwave oscillators are of very low magnitude compared with the carrier power. It follows that the a.m. and f.m. noise modulation indices are very small ($\ll 1$), which is very fortunate as it means that only the first-order sidebands of the f.m. spectrum need be considered. Thus in a simple diode detector the a.m. noise is demodulated to give an output at the modulation frequency, whereas the f.m. noise produces practically no demodulated output. Because the modulation index is small the noise measuring technique may take one of the following forms.

The output of a microwave oscillator under test may be fed into the crystal detector of a direct-detection system which in the simplest case consists of the detector followed by a tunable amplifier and an indicating device. The upper and lower a.m. noise sidebands beat with the carrier to produce an output and can be measured, whereas the beat products of carrier and f.m. sidebands are in antiphase and cancel. The beat products between carrier and background noise also produce an output at the detector so that this method indicates both a.m. and background noise.

In order to measure f.m. noise in a direct-detection system, some form of f.m. to a.m. conversion (and preferably the reverse process as well) is necessary. Such a conversion occurs if the phase of the carrier is shifted by $\frac{1}{2}\pi$ with respect to the sidebands. In practice this can be achieved by suppressing the carrier with a suitable filter and re-introducing it with the required phase shift. The phase of the sidebands is also changed by the suppression filter, but since the upper sidebands are phase-retarded and the lower sidebands phase-advanced by the same amount with respect to the original carrier, it may be shown that the demodulated output is unaffected. This method is not suitable for measurements close to the carrier frequency because of the finite bandwidth of the carrier suppression filter.

The conversion can also be achieved using a microwave discriminator. With a push-pull arrangement cancellation of the a.m. components is possible. As the bandwidth of a microwave discriminator is restricted, this conversion method can be used only for measurements relatively near the carrier. Thus the two conversion methods are complementary and enable a wide frequency range to be covered. It should be noted that again the background noise component is included in the measurement.

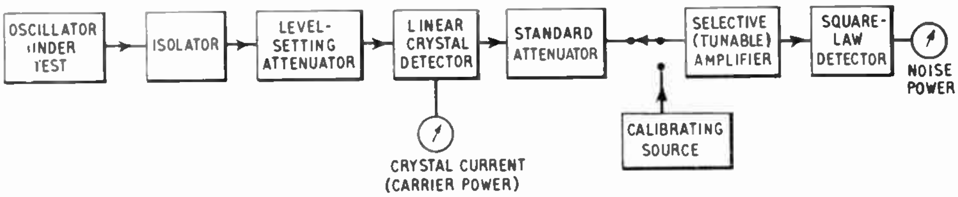


Fig. 1. Direct-detection measurement of a.m. + background noise spectra up to 100 Mc/s.

It has been shown that the direct-detection method can be adapted to measure a.m. plus background noise and f.m. plus background noise. On the other hand, with a superheterodyne system the total noise can be measured. The use of a superhet method for oscillator noise measurements presents certain difficulties. For example, in a direct-detection system the carrier itself performs the function of the local-oscillator drive and is converted to d.c. The superhet, on the other hand, has its own local oscillator and in the mixer any applied signal, which includes the carrier, must be well below the applied local-oscillator level. Thus for maximum sensitivity the carrier must be removed. As before, and because of the finite bandwidth of the carrier suppression filter, it is not possible to make measurements close to the carrier. At X-band the limit is a few hundred kilocycles per second.

Whatever system is used, noise spectrum measurements require the incorporation of a narrow-band tunable amplifier to select and amplify noise samples over the required range. If only r.m.s. values are required a wide-band amplifier is more convenient. Consideration must be given to the relationship between the bandwidth, B, of this amplifier and the combined time-constant, T, of the second (square-law) detector and indicating instrument. The probable error of a single reading is proportional to $(2BT)^{-\frac{1}{2}}$ so that for high resolution T must be large.[7]

It is preferable for the measured noise to be compared directly with that of a standard microwave fluorescent noise source as a knowledge of the detector (mixer) and amplifier characteristics is not then required. However, for frequencies close to the carrier in a direct-detection system, the noise powers encountered may exceed those of available microwave noise sources and it may become necessary to use either a noise diode or a sinusoidal signal as the reference power. In this case, a detailed knowledge of the detector, or detector plus amplifier, characteristics is required.

3. Direct-Detection Methods

In a simple direct-detection (crystal-video) microwave receiver[8] the incoming r.f. signal is demodulated in the crystal detector, and the resulting output is amplified in a high-gain video amplifier. The sensitivity of such a receiver is poor mainly because the detector works at a low level. For oscillator noise measurements the detector can work at a high level as sufficient carrier power is available from the oscillator under test and the sensitivity is higher. If the applied crystal drive is made the same as is customary in a crystal mixer, the sensitivity will be similar to that of a crystal-mixer superhet. Thus providing the combination of detector and first amplifier stage is suitably designed it is possible to measure to within a few decibels of the thermal noise level (-174 dB/c/s), in other words carrier/noise ratios of ~ 170 dB/c/s. The carrier may be thought of as supplying the local-oscillator power, the sidebands being the signal. The processes in a detector and in a mixer are, of course, essentially the same. The relatively high carrier power is converted to d.c. and cannot cause undesired effects in the succeeding amplifier, as may occur in the superhet system with inadequate suppression.

3.1. *Measurement of A.M. plus Background Noise*

3.1.1. I.f. calibration method

The general lay-out of a system for the direct-detection measurement of a.m. plus background noise spectra at frequencies from a few cycles per second upwards to, if necessary, 100 Mc/s and more, is given in Fig. 1. The oscillator output is passed through an isolator to eliminate any undesirable effects due to reflections, and a level-setting attenuator, and then enters a linear crystal detector which is followed by a standard attenuator, an i.f. amplifier to scan the spectrum, and a power indicator. The amplifier must be calibrated either with a noise diode or with a sinusoidal signal generator; in the latter case the amplifier bandwidth must also be accurately known. The result is conveniently expressed as a carrier/noise power ratio, and may be determined from the measured noise power, and the crystal current reading which is a measure of the carrier power. Normally the carrier power is related to the noise power contained in a band of width 1 c/s. As crystal noise also contributes to the indicated noise power this has to be measured separately for a correction to be applied.

Several systems based on this principle have been reported,[9–13] most of which employ a wide-band amplifier to measure the average carrier/noise ratio.

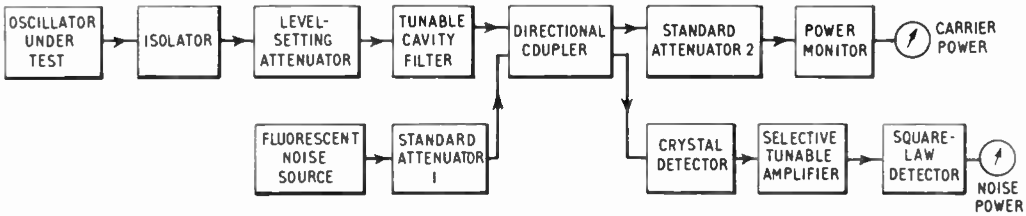


Fig. 2. Cavity-filter noise measuring circuit.

Ashby and Dyott[14] describe the use of a coaxial vacuum diode, instead of a crystal detector, which may be of advantage when measuring noise in high-power devices.

Although the circuit of Fig. 1 can be used for frequencies close to the carrier, the main interest in the a.m. plus background noise spectrum is usually at higher sideband frequencies in connection with local-oscillator applications. At these higher frequencies the oscillator noise level is normally sufficiently low that a microwave noise standard in front of the crystal detector may be used as a reference noise power. This has the great advantage that the detector need not be linear and no knowledge of the crystal-noise contribution, the detector efficiency, or the amplifier characteristics, is necessary.

Various methods which use a fluorescent noise source to provide the reference noise power are described below. They are substitution methods in which firstly the noise from the oscillator under test produces a reading on the output meter. Secondly, the oscillator noise is removed to obtain a "clean" carrier, and a known amount of noise from the fluorescent noise source is then added to produce the same reading. The methods differ mainly in the way the clean carrier is produced.

3.1.2. Cavity-filter method

The oscillator noise sidebands can be suppressed with a bandpass filter, which in its simplest form may be a single transmission cavity tuned to the carrier frequency. A filter which does not require the physical insertion and removal of components, or some form of microwave switching, has been described by Court.[15]

Figure 2 shows a block diagram of the cavity-filter noise measuring circuit. A directional coupler is used to connect both the oscillator and the noise source branches to the crystal detector. The measuring procedure consists of obtaining identical readings on the output meter (*a*) with the noise source switched off and the filter in the "out" condition, and (*b*) with the noise source on and the filter inserted. By adjusting the level-setting attenuator the power incident on the crystal must be kept constant for both cases.

The indication of the output meter is proportional to:

$$\text{case } (a) \quad N_1 = \left(\frac{N_o}{L_c} + N_c + N_a\right) G_a \qquad \ldots\ldots(1)$$

$$\text{case } (b) \quad N_2 = \left(\frac{N_o}{S_f L_c} + \frac{N_n}{L_d L_c A_1} + N_c + N_a\right) G_a \quad \ldots(2)$$

where N_o is the oscillator noise at the mixer input

- N_c is the noise power contributed by the crystal (referred to detector output)
- N_n is the noise power of the fluorescent noise source
- N_a is the noise power contributed by the amplifier (referred to amplifier input)
- L_d is the loss of noise-source power in the directional coupler ($L_d > 1$)
- L_c is the conversion loss of the crystal detector ($L_c > 1$)
- S_f is the sideband suppression factor of the filter ($S_f \geqslant 1$)
- G_a is the amplifier gain
- A_1 is the attenuation factor of the standard attenuator ($A_1 \geqslant 1$)

and where the noise powers are those within the amplifier bandwidth. Thus when $N_1 = N_2$

$$N_o = \frac{N_n S_f}{L_d A_1 (S_f - 1)} \qquad \ldots\ldots(3)$$

To determine the carrier/noise ratio, the carrier power must also be measured and referred to the detector input. The frequency range of interest is covered by tuning the amplifier.

This method has the advantage of simplicity but some loss of signal strength is inevitable because of the insertion loss of the cavity. Furthermore because of the finite cavity bandwidth, measurements cannot be made at frequencies close to that of the carrier.

3.1.3. Balanced-mixer method

This method has been suggested by Smith.[16] As the name implies, a balanced-mixer[17] † is used to

† Although we are dealing with a direct-*detection* system the term balanced-*mixer* is used as the device is normally known by this name.

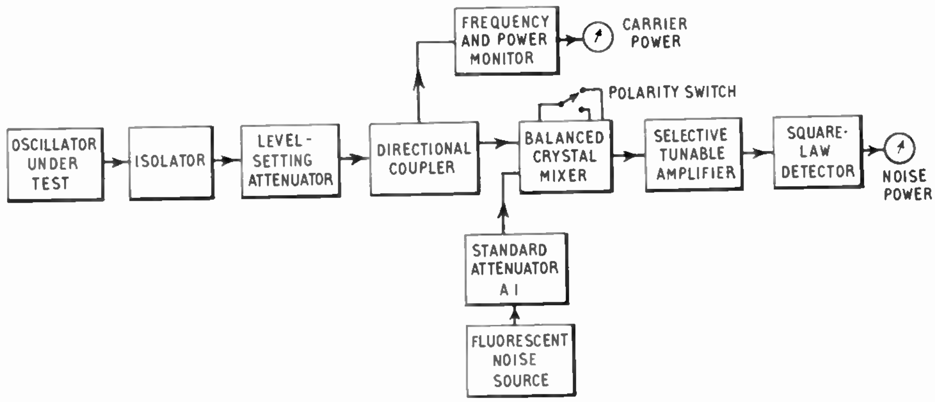


Fig. 3. Balanced-mixer noise measuring circuit.

suppress the oscillator noise sidebands by a known amount. Both terminals of one of the mixer crystals are isolated, so that the polarity of its output may be reversed in order to produce both balanced and unbalanced conditions. Care must be taken to ensure that this reversal does not cause any significant change in the impedance presented to the following amplifier.

A block diagram of the balanced-mixer noise measuring circuit is given in Fig. 3. The oscillator under test is connected to one input arm of the mixer and the fluorescent noise source to the other. Identical readings are obtained on the output meter (*a*) with the noise source switched off and the mixer unbalanced, and (*b*) with the noise source on and the mixer balanced. The noise power is evaluated in a similar manner to that outlined for the cavity-filter method and the oscillator noise power at the mixer input is

$$N_o = \frac{N_n S_m}{A_1(S_m - 1)} \qquad \ldots\ldots(4)$$

where S_m is the sideband suppression factor of the balanced mixer ($S_m > 1$).

The suppression factor of the mixer must be known and can be found by applying to the mixer a microwave signal, amplitude modulated at the required sideband frequency, and measuring the output in the balanced and the unbalanced conditions. A simple crystal modulator may be used. To ensure accuracy the modulation sidebands should be well above the level of the noise which the signal might introduce.

The balanced-mixer method has the advantage that there is no additional loss of signal strength, as occurs in the cavity-filter method due to the insertion loss of the filter. In addition the suppression in the balanced mixer is effective down to the smallest frequency separations from the carrier. On the other hand balanced amplifier input circuits are required.

3.1.4. Measurement of high-level noise with a fluorescent noise source

With the two methods just described the oscillator noise is measurable as long as its level does not, at the detector (mixer) input, exceed that of the fluorescent noise source which is 15 to 18 dB above thermal noise. Noise in O-type oscillators† over the i.f. range 15 to 60 Mc/s seldom exceeds this limit. However the lower sideband range is often of interest, and here, particularly in M-type tubes, the noise generally exceeds that of the fluorescent noise source. The oscillator noise, together with the carrier, cannot simply be attenuated in front of the detector as then the correct "local-oscillator" drive at the crystal may no longer be maintained. However an extension of the measuring range is made possible by the insertion of an attenuator between the detector and the amplifier to reduce the oscillator noise. During the reference measurement with the noise source this attenuator is switched out.

A brief analysis of the balanced-mixer method modified in this manner is given in the following, where the additional attenuator A2, inserted behind the mixer, has an attenuation factor A_2 and the attenuator A1 is set to zero attenuation. A reading is first taken of the suppressed oscillator, plus calibration source, noise. The mixer is then unbalanced, the noise source switched off and A2 adjusted to give the same reading on the output meter. The oscillator noise at the mixer input terminals, previously given by eqn. (4), then becomes

$$N_o = \frac{A_2(N_n + N_c L_c) - N_c L_c}{1 - \dfrac{A_2}{S_m}} \qquad \ldots\ldots(5)$$

† In O-type devices (e.g. klystrons, travelling-wave tubes) there is no steady transverse field, whereas in M-type, or crossed-field, devices (e.g. magnetrons) steady transverse electric and magnetic fields are present. See C. H. Dix and W. E. Willshaw "Microwave valves", *J.Brit.I.R.E.*, **20**, p. 577, August 1960.

The mixer properties must now be known. Conversion loss and crystal noise can be measured by standard methods[18, 19] although it is possible to determine $N_c L_c = N_c'$, the crystal noise referred to the mixer input, simply by an additional measurement. This is done by equalizing output readings of (*a*) suppressed oscillator noise and (*b*) attenuated and suppressed oscillator noise plus calibration source noise. This gives

$$\frac{N_o}{S_m} + N_c' = \frac{\frac{N_o}{S_m} + N_c' + N_n}{A_2'} \qquad \ldots\ldots(6)$$

hence

$$N_c' = \frac{\frac{N_o}{S_m}(1 - A_2') + N_n}{A_2' - 1} \qquad \ldots\ldots(7)$$

This value can be inserted in eqn. (5), but a higher accuracy is achieved if eqn. (6) is combined with another expression obtained by equating the suppressed oscillator noise and the attenuated unsuppressed oscillator noise, namely

$$\frac{N_o}{S_m} + N_c' = A_2''(N_o + N_c') \qquad \ldots\ldots(8)$$

Thus the measurements described by eqns. (6) and (8) give the oscillator noise as

$$N_o = \frac{(A_2'' - 1) S_m N_n}{(A_2' - 1)(S_m - 1)} \qquad \ldots\ldots(9)$$

The range extension is greatest for small values of S_m and N_c. The limit also depends to a large degree on the accuracy with which A_2 and S_m are known. However, it is not possible to achieve an extension of the measurable noise level by more than about 20 dB in this way.[20] This constitutes a marked improvement over the simple direct-detection systems, but is, unfortunately, not always sufficient for measurements at sideband frequencies down to a few hundred kilocycles per second.

Since the oscillator carrier must provide a given optimum "local oscillator" power level at the crystal detector, normally about 1 mW, a further extension of the measuring range towards higher noise levels may be achieved by attenuating the noise sidebands independently of the carrier. A block diagram of a system for separating the noise sidebands and the carrier is shown in Fig. 4. A narrow band (< 1 Mc/s) carrier suppression (c.s.) filter removes the carrier and thus only the noise sidebands reach input arm 1 of the balanced mixer. Some of the oscillator output is diverted from the c.s. filter by means of a directional coupler and is supplied to arm 2 of the mixer through an auxiliary branch. The balanced mixer is adjusted to suppress noise sidebands entering arm 2 with the carrier, thus providing the necessary "clean" carrier. The phase of this re-introduced carrier is adjusted by means of the phase shifter in the auxiliary branch, so that it differs by $\pm n\pi$ ($n = 0, 1, 2 \ldots$) from that of the suppressed carrier component associated with the noise sidebands.

When equal output-meter readings are obtained with (*a*) the oscillator noise, and (*b*) the fluorescent noise source, connected to the mixer arm 1 by means of waveguide switch 2, then the oscillator noise power at the reference plane P is

$$N_o = \frac{N_n A_1}{A_2} \qquad \ldots\ldots(10)$$

Depending on the oscillator noise level, either standard attenuator A_1 or A_2 is normally set to unity attenuation. The carrier power at plane P can be determined by connecting the power monitor via waveguide switch 1, and detuning the cavity in the c.s. filter so

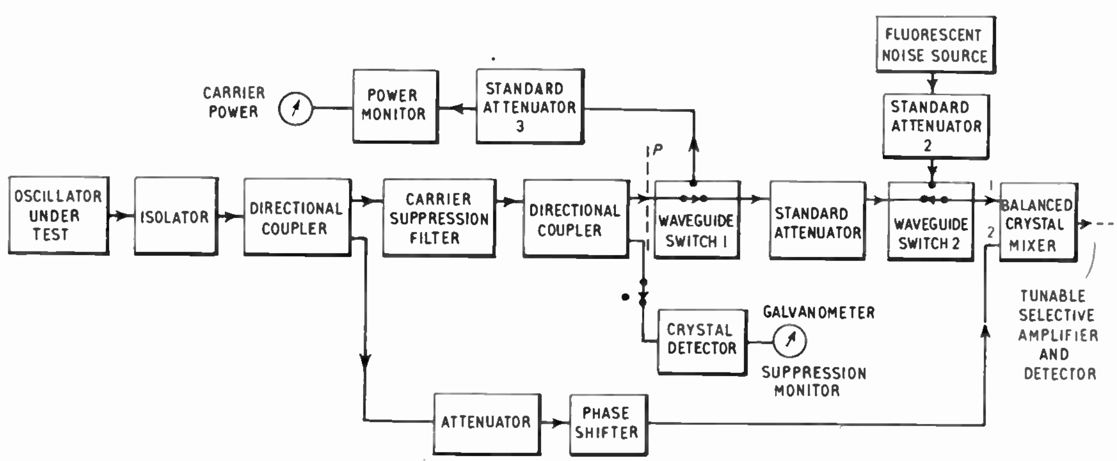


Fig. 4. System for separating noise sidebands from the carrier.

that the carrier is subjected only to the filter insertion loss. During this measurement the detector crystal in the suppression monitor branch must, naturally, be protected from overloading. Due to the finite bandwidth of the filter, the measured noise power has to be corrected at low sideband frequencies according to the suppression-frequency curve of the filter.

The range extension which may be achieved depends on the balanced mixer suppression factor, S_m. The authors have found that with a matched pair of crystals and with suitable adjustment of the crystal drive and output impedance separately for each crystal, a suppression of 50 dB over a wide frequency range is possible. Thus if the residual noise with the re-introduced carrier after suppression is to be at least 8 dB, say, below the noise to be measured then, for a carrier power of 1 mW, carrier/noise ratios as high as 116 dB/c/s can be measured. This assumes a calibration noise level of 15·5 dB at mixer input 1. The procedure for setting the phase shifter to provide the required phase of the "clean" carrier depends on the design of the carrier suppression filter and is usually quite simple. It is necessary to correct for the variable phase change introduced by attenuator A1.

The upper limit of noise measurement can also be extended by amplifying the output from a fluorescent noise source using a travelling-wave tube in order to produce a high-level comparison source.[21] This would not seem to be a very satisfactory process since the gain and the noise level of the travelling-wave tube have to be accurately known and must be kept constant. Furthermore the amount of range extension is less than that afforded by the preceding methods unless two travelling-wave tubes are used, when the difficulties and the expense become considerable.

3.2. *Measurement of F.M. plus Background Noise*

Frequency-modulation noise can be measured by a direct-detection process if it is first converted to amplitude-modulation noise, providing the original a.m. components do not contribute to the measurement or that a suitable correction can be made. The background noise is also measured since it is completely uncorrelated with the carrier, but at low sideband frequencies it is negligible compared with the modulation noise.

3.2.1. Change of carrier phase

For single-tone modulation it may be shown that if the carrier is phase advanced or retarded by $\frac{1}{2}\pi$, amplitude modulation is converted to frequency modulation, and vice versa. The same consideration applies to the continuous low-order modulation noise spectrum of microwave oscillators, and thus by shifting the carrier phase by $\frac{1}{2}\pi$ a direct-detection system can be used to measure f.m. noise quite independently of the a.m. noise.

In practice a conversion of this kind may be achieved by suppressing the carrier and re-introducing it with the required phase shift, using the circuit of Fig. 4 which has been described above.

The measuring procedure for determining the f.m. plus background noise power spectrum is the same as that described in Section 3.1.4. with the exception, of course, of the carrier phase change. With this method it is possible to measure f.m. noise spectra up to the same level as for a.m. spectra.

3.2.2. Use of microwave discriminator

If, in the simple direct-detection system shown in Fig. 1, a microwave discriminator is inserted in front of the linear crystal detector, frequency fluctuations are converted into amplitude fluctuations and thus become measurable. The amplifier following the detector can either be selective for plotting the carrier frequency deviation spectrum, or it can have a wide bandwidth to measure the r.m.s. carrier frequency deviation as represented by the f.m. noise power within this bandwidth. The spectrum range which can be covered without introducing inaccuracies is limited to the linear portion of the discriminator characteristic. Microwave discriminators make use of cavities and thus at X-band, for example, it is difficult to measure beyond about 10 Mc/s. For modulation frequencies in this range background noise is usually negligible, and the crystal noise contribution is small. The calibration has to be undertaken behind the detector with a noise diode or a sinusoidal signal generator.

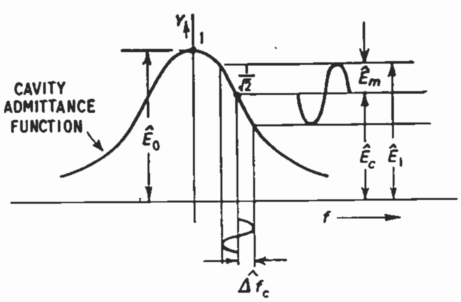


Fig. 5. Cavity admittance function with a frequency-modulated signal.

A simple form of discriminator is obtained by the use of a transmission cavity tuned to one of its half-power points so that slope conversion occurs at this point of maximum slope and linearity. Several measuring systems based on this principle have been reported.[1, 2, 11, 12, 22] An expression for the r.m.s. carrier frequency deviation† can be derived using the following simple physical considerations.

† Defined as the deviation which would be caused by a sinusoidal modulation producing the same sideband power as measured here.

Consider Fig. 5 which shows the cavity admittance function with a frequency-modulated input signal, the mean frequency of which is at one of the cavity half-power frequencies. Because of the slope of the cavity characteristic the output from the cavity is amplitude modulated.

The mean signal amplitude at the cavity output is

$$\hat{E}_c = \hat{E}_o |Y|_h = \frac{\hat{E}_o}{\sqrt{2}} \qquad \ldots\ldots(11)$$

where $\hat{E}_o$ is the peak amplitude of the output signal with no frequency modulation and with the cavity tuned to the signal frequency, and $|Y|_h$ is the magnitude of the normalized cavity admittance function at the half-power points.

The peak amplitude of the output signal is

$$\hat{E}_1 = \frac{\hat{E}_o}{\sqrt{2}} + \hat{E}_o|\delta Y| = \frac{\hat{E}_o}{\sqrt{2}} + \hat{E}_o \widehat{\Delta f_c} \left|\frac{\mathrm{d}Y}{\mathrm{d}f}\right|_h \ldots\ldots(12)$$

Here $\widehat{\Delta f_c}$ is the peak frequency deviation of the input signal. Hence the peak variation in amplitude of the output signal is given by

$$\hat{E}_m = \hat{E}_1 - \hat{E}_c = \hat{E}_o \widehat{\Delta f_c} \left|\frac{\mathrm{d}Y}{\mathrm{d}f}\right|_h \qquad \ldots\ldots(13)$$

and thus

$$\widehat{\Delta f_c} = \frac{\hat{E}_m}{\sqrt{2}\hat{E}_c \left|\frac{\mathrm{d}Y}{\mathrm{d}f}\right|_h} \qquad \ldots\ldots(14)$$

The subsequent analysis has to be carried out assuming an envelope detector is used, since the time constant, RC, of the detector output circuit is normally large compared with the periodic time at the microwave signal frequency.

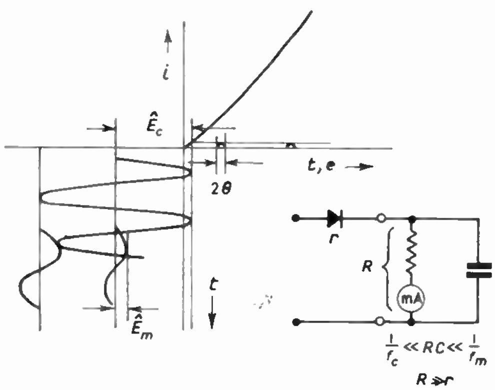


Fig. 6. Calculations of carrier frequency deviation using an envelope detector.

Under the assumption that the resistive part of the output impedance, R, is very large compared with the forward resistance, r, of the diode, and that the time constant is small compared with the periodic time at the modulation frequency, then with the notation shown in Fig. 6 the d.c. component of the detector output voltage is very nearly

$$E_{\mathrm{dc}} = \hat{E}_c \qquad \ldots\ldots(15)$$

and hence the d.c. power

$$P_{\mathrm{dc}} = \frac{\hat{E}_c^2}{R} \qquad \ldots\ldots(16)$$

The amplitude of the voltage variation due to the modulation of the input wave is $\hat{E}_m$, and the a.c. power is

$$P_{\mathrm{ac}} = \frac{1}{2}\frac{\hat{E}_m^2}{R} \qquad \ldots\ldots(17)$$

Combining eqns. (16) and (17) yields

$$\frac{\hat{E}_m}{\hat{E}_c} = \sqrt{\frac{2P_{\mathrm{ac}}}{P_{\mathrm{dc}}}} \qquad \ldots\ldots(18)$$

which inserted in eqn. (14) gives the r.m.s. carrier frequency deviation:

$$(\Delta f_c)_{\mathrm{r.m.s.}} = \sqrt{\frac{P_{\mathrm{ac}}}{2P_{\mathrm{dc}}}} \left\{\left|\frac{\mathrm{d}Y}{\mathrm{d}f}\right|_h\right\}^{-1} \qquad \ldots\ldots(19)$$

With the help of this expression the carrier frequency deviation can be calculated from the measured carrier power (P_{dc}) and the noise power (P_{ac}).

For the assumed condition of $R \gg r$ the conduction angle, θ, is almost zero and the analysis is valid even if the diode characteristic is not linear. On the other hand, when $R \simeq r$, as is often the case, the conduction angle is not very small, but if the diode characteristic is linear eqns. (16) and (17) become

$$P_{\mathrm{dc}} = \frac{(\hat{E}_c \cos\theta)^2}{R} \quad \text{and} \quad P_{\mathrm{ac}} = \frac{1}{2}\frac{(\hat{E}_m \cos\theta)^2}{R} \qquad \ldots\ldots(16a, 17a)$$

and the expression for the r.m.s. carrier frequency deviation remains unchanged. In practice the characteristic is markedly curved near the origin and it is necessary to apply a sufficiently large input signal that the diode operates mainly over the linear part of the characteristic.

In eqn. (19) P_{ac} is assumed to be entirely due to f.m. noise but, since no limiting action takes place in the discriminator, the a.m. noise component, modified by the cavity response curve, also appears at the detector. The f.m. and a.m. noise contributions are usually partially correlated so that different noise powers are measured at the two half-power points of the cavity. At one half-power point the vector sum, and at the other the difference, of the correlated f.m. and a.m. components is obtained, since the relative phase of the a.m. and f.m. components changes by π between these points. The difference in measured noise power can be quite marked even with a relatively small a.m. component.

Figure 7 shows curves of measured noise power and crystal current for a CO43 X-band backward-wave oscillator, obtained when the discriminator cavity was tuned about the carrier frequency. The amplifier bandwidth was about 4 Mc/s. The peaks occur at the half-power points, and the fact that the higher peak occurs at the upper half-power point agrees with the negative frequency-pushing effect† normally observed in backward-wave oscillators. At the upper half-power point beam current fluctuations result in two in-phase components of the amplitude variation at the cavity output, one component due to amplitude, the other due to frequency variations. At the lower half-power point these two resulting components are in antiphase.

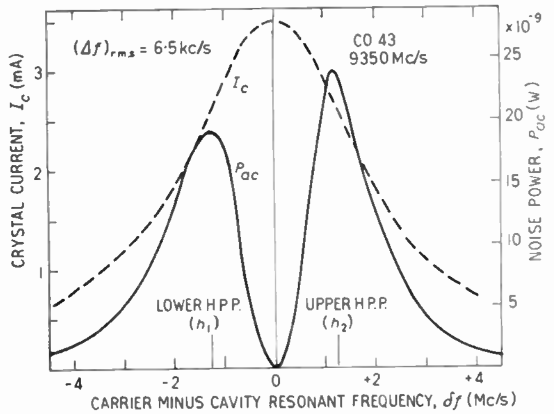


Fig. 7. Measured noise power and crystal current curves for a CO43 X-band backward-wave oscillator.

In order to calculate the carrier frequency deviation with the help of eqn. (19), it is necessary, therefore, to determine that part of the noise power which is due solely to the frequency fluctuations. This can be done from three noise measurements, one at the upper half-power point, one at the lower half-power point, and a third measurement without the cavity discriminator to determine the a.m. noise. The latter reading has to be corrected for the modification of this noise by the cavity response. If amplitude fluctuations are represented by A and frequency fluctuations by F, the noise power measured at one half-power point is proportional to $\overline{(A+F)^2}$, that measured at the other half-power point is proportional to $\overline{(A-F)^2}$, whereas the a.m. noise power is proportional to $\overline{A^2}$. From these three values the frequency fluctuation component becomes

$$\overline{F^2} = \frac{\overline{(A+F)^2} + \overline{(A-F)^2}}{2} - \overline{A^2}$$

and so one obtains for the f.m. noise power

$$(P_{ac})_{fm} = \tfrac{1}{2}[(P_{ac})_{h_1} + (P_{ac})_{h_2}] - P_{am} \quad \ldots\ldots(20)$$

where $(P_{ac})_{h_{1(2)}}$ is the measured noise power at the lower (upper) half-power point, and P_{am} is the a.m. noise power.

In general, the effect of the a.m. noise is less marked the higher the Q-factor of the cavity because of the higher conversion efficiency, but at the expense of a restricted bandwidth. If the a.m. component is not negligible it is advantageous to use a push-pull discriminator which, in addition, has an increased linear range. A push-pull discriminator found to be very suitable, since it requires only a single cavity, has been described by Colani.[30]

A further unwanted noise contribution is that caused by the detector crystal. If necessary a correction may be applied but for measurements of r.m.s. carrier frequency deviation the crystal noise contribution is usually negligible.

4. The Superheterodyne Method

A microwave superheterodyne receiver can also be used to measure oscillator noise. Such a receiver inherently measures the upper and lower noise sidebands separately, and for most practical purposes the *total* noise—a.m., f.m., and background components—contained in the spectrum sample is determined. However any correlated a.m. and f.m. components do not add up in an r.m.s. manner. Theoretically, in the most unfavourable case with full correlation, a.m. and f.m. components of equal magnitude, and a phase difference of $\frac{1}{2}\pi$ between them, then the modulation components add in one sideband and cancel in the other. Fortunately, there seldom appears to be a marked time lag between amplitude and frequency modulation, and over most of the spectrum range one of the modulation components usually predominates.

In order to ensure correct superheterodyne action, the signal power level at the mixer, which here includes the oscillator carrier, should be at least 10 dB below that of the local oscillator. If an attenuator is used to reduce the carrier power at the mixer to the necessary level, the noise sidebands are at the same time attenuated so much that they may no longer be detectable. Thus it is preferable to remove the carrier, with a narrow-band carrier-suppression filter, so that apart from the insertion loss the full sideband output of the oscillator under test may be applied to the measuring system.

Dalman and Rhoads[23] used a superhet system with carrier suppression. However their filter was rather

† *Negative* frequency-pushing implies that with increased beam current the frequency *decreases* whereas the output power increases.

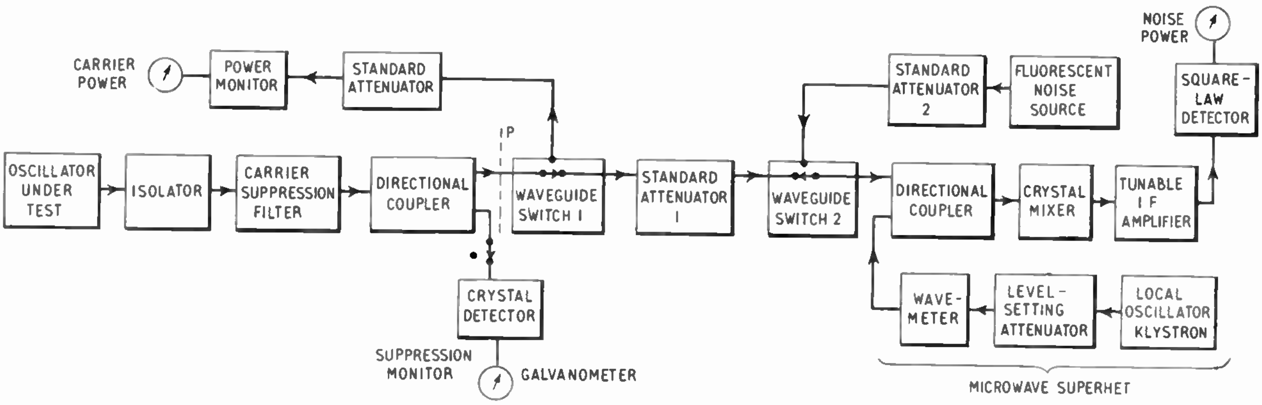


Fig. 8. Microwave superheterodyne system for noise spectrum measurements.

crude and of such a wide bandwidth that even at frequency separations of 50 Mc/s from the carrier the attenuation was still as high as 15 dB so that large corrections were necessary. Furthermore these authors assumed that the system measured local-oscillator noise contributions whereas their results included the f.m. noise component which predominates at the lower sideband frequencies, and does not of course contribute to local-oscillator noise.

A block diagram of a microwave superhet system for noise spectrum measurements which has been developed by the present authors is shown in Fig. 8. The oscillator output is passed through the narrow-band ($<$ 1 Mc/s) c.s. filter to the crystal mixer. A v.h.f. communications receiver is used as a selective, tunable, i.f. amplifier, thus making it essentially a double-superhet system. The frequency of the klystron local oscillator is kept constant and tuning is performed with the v.h.f. receiver. Measurements are carried out in a similar way to that described in Section 3.1 for a.m. plus background noise. First the fluorescent noise source as the reference, and then the oscillator, are connected to the mixer and attenuators 1 and 2, respectively, adjusted to give equal readings on the output meter, whence

$$N_o = \frac{N_n A_1}{A_2}$$

where N_o is the oscillator noise at the reference plane P,

N_n is the fluorescent noise source power, and

A is the attenuation factor of the attenuators ($A \geqslant 1$).

To express the result as a carrier/noise ratio, the carrier power is measured in the power monitor after detuning the c.s. filter, and the noise power corrected where the filter suppression exceeds the insertion loss. With this method there is no upper limit to the noise level which can be measured.

The factors governing the choice of local oscillator frequency and first i.f. tuning range are illustrated by reference to a system designed to cover the spectrum range $\pm$ 0·5–114 Mc/s from the carrier at X-band. As illustrated in Fig. 9, the l.o. frequency is higher (lower) than the test-oscillator carrier frequency for measurement of the upper (lower) sideband. A v.h.f. receiver is employed in the i.f. part of the microwave superhet in order to minimize the possibility of image reception and to allow the use of a single-ended mixer, since noise contributions from the l.o. klystron are normally negligible at very high frequencies.

Measurements can be made, at X-band, to within about 200 kc/s of the carrier. This limit is set mainly by the relative frequency drift of the test and local oscillators and may be achieved by using high-stability power supplies for both oscillators. Relative frequency drift can be avoided by "locking" the two oscillators,[24–26] the local oscillator then follows the drifting test oscillator. For the particular frequencies shown in Fig. 9 the local oscillator must be locked to a frequency 164 Mc/s away from the test-oscillator carrier. This may be done by supplying the difference frequency from a stable v.h.f. signal generator to a

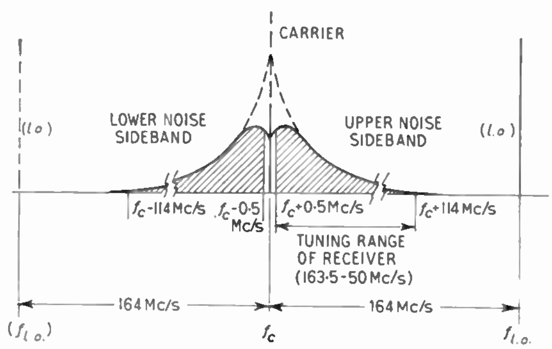


Fig. 9. Spectrum of carrier and noise sidebands.

microwave modulator, to which is also fed the output from the test oscillator. The appropriate difference frequency is then used to lock the local oscillator. Sufficient isolation has to be provided between test and local oscillator to ensure that locking takes place only in the desired direction. One of the properties of frequency synchronization is that the locked oscillator tends to assume the noise spectrum of the locking source. This may lead to increased l.o. noise, necessitating the use of a balanced mixer in order to prevent a reduction in the sensitivity of the receiver. Locking has been used successfully but requires a large amount of additional equipment.

5. Measurement of Correlation

An important quantity in oscillator noise investigations is the cross-correlation between a.m. and f.m. noise. Correlation is to be expected since both the amplitude and the frequency of oscillation are influenced by certain factors, for example both the output power and the frequency depend on the beam current. This is also true of microscopic fluctuations and beam noise is one of the factors producing correlation between a.m. and f.m. noise. At the same time there is a strong correlation between beam and modulation noise.

In the measurement of carrier frequency deviation employing a transmission cavity (Sect. 3.2.2), the curve obtained for the noise power, P_{ac}, as a function of cavity tuning, e.g. Fig. 7, has two peaks of unequal height at the cavity half-power points, due to correlation. The relative height of the peaks can be used to determine the correlation coefficient.

In general the degree of correlation is expressed by the cross-correlation coefficient, c, which for two fluctuating quantities x and y is defined[7] as

$$c = \frac{\overline{xy}}{\sqrt{\overline{x^2}}\sqrt{\overline{y^2}}} \qquad (-1 \leqslant c \leqslant 1) \qquad \ldots\ldots(21)$$

In order to relate the cross-correlation coefficient to the quantities measured with the transmission-cavity system, eqn. (21) has to be rewritten in the form

$$c = \frac{\overline{(x+y)^2} - \overline{x^2} - \overline{y^2}}{2\sqrt{\overline{x^2}\,\overline{y^2}}} \qquad \ldots\ldots(21a)$$

and finally

$$c = \frac{\overline{(x+y)^2} - \overline{(x-y)^2}}{4\sqrt{\overline{x^2}[\frac{1}{2}\overline{(x+y)^2} + \frac{1}{2}\overline{(x-y)^2} - \overline{x^2}]}} \qquad \ldots\ldots(21b)$$

It has been shown above that if x represents the amplitude fluctuations and y the frequency fluctuations, $\overline{(x + y)^2}$ is proportional to the measured noise power at one cavity half-power point, $\overline{(x - y)^2}$ proportional to that at the other half-power point, and $\overline{x^2}$ is proportional to the measured a.m. noise power. The correlation factor, thus, becomes

$$c = \frac{(P_{ac})_{h_1} - (P_{ac})_{h_2}}{4\sqrt{P_{am}[\frac{1}{2}(P_{ac})_{h_1} + \frac{1}{2}(P_{ac})_{h_2} - P_{am}]}} \ldots\ldots(21c)$$

It should be noted that with the above definition the correlation coefficient is lower when a time lag exists between the amplitude and frequency fluctuation components which are caused by the same generating event. In addition, any background noise present will tend to reduce the calculated correlation value although the effect of background noise normally is negligible.

A second method of finding the degree of cross-correlation is to produce "scatter-plots" (random Lissajous figures) on an oscillograph.[12, 27, 28] Two direct-detection systems are required to supply a.m. noise to one pair of plates and f.m. noise to the other pair of plates (producing a deflection perpendicular to the first) of the oscillograph tube. The f.m. channel must be free from a.m. components so that, unless the phase-shift method is used, a push-pull discriminator is necessary. A pulse generator serves to modulate the beam intensity of the oscillograph tube so that points of light (at the desired sampling rate) are produced on the screen. If both channels have the same bandwidth and phase response, and if the two components applied to the electrodes are adjusted to be of the same magnitude, then the light dots will form a "filled-in" ellipse, the extremes being a circle for zero correlation and a line with a slope of 45 deg for full correlation. In general, the cross-correlation coefficient is

$$c = \frac{1-p^2}{1+p^2}$$

p being the ratio of the minor and major axes of the ellipse. The brightness of the ellipse diminishes with increasing distance from the centre since the number of light dots decreases. For all practical purposes, however, it does not appear to be essential to modulate the beam intensity. Instead the moving spot may be observed giving the added advantage that the area of the ellipse appears more clearly defined. The error in estimating the correlation coefficient in this way[28] does not exceed $\pm$ 0·1.

The scatter-plot method can also be used to measure the correlation between the (low-frequency) oscillator beam noise and either a.m. or f.m. noise. A sample of the beam noise is taken from a small resistor placed in the supply line and applied to one channel of the oscillograph, and the modulation noise is applied to the other channel.

6. Discussion

As shown above there are three distinct types of system which may be used for the measurement of oscillator noise, namely the a.m. and f.m. direct-detection systems and the superhet system. Each of these responds to more than one noise component and the question arises whether it is possible to separate these various components. Since there are three components and three distinct types of measurement then it might be thought that a complete separation of oscillator noise into its a.m., f.m. and background components is possible. That this is not the case may be shown as follows. As before the instantaneous values of the fluctuation components at the upper or lower sideband frequencies due to amplitude modulation may be described by A, and those due to frequency modulation by F. Furthermore it will be assumed that the background noise is symmetrical about the carrier frequency and that the instantaneous background fluctuations at the particular sideband frequency investigated is B. Then the noise powers associated with the individual a.m., f.m., and background, components are, respectively

$$N_{\mathrm{am}} = k\overline{A^2} \qquad \ldots\ldots(21)$$

$$N_{\mathrm{fm}} = k\overline{F^2} \qquad \ldots\ldots(22)$$

$$N_{\mathrm{b}} = k\overline{B^2} \qquad \ldots\ldots(23)$$

where k is a constant.

In a.m. direct-detection measuring systems the noise components arising from the upper and lower a.m. sidebands are in phase, whereas the two background components are uncorrelated. Hence the noise power indicated is

$$N_{d(\mathrm{am})} = k[\overline{(2A)^2} + \overline{2B^2}]$$

or with eqns. (21) and (23)

$$N_{d(\mathrm{am})} = 4N_{\mathrm{am}} + 2N_{\mathrm{b}} \qquad \ldots\ldots(24)$$

Similarly f.m. direct-detection systems measure a power

$$N_{d(\mathrm{fm})} = k[\overline{(2F)^2} + \overline{2B^2}]$$

or

$$N_{d(\mathrm{fm})} = 4N_{\mathrm{fm}} + 2N_{\mathrm{b}} \qquad \ldots\ldots(25)$$

Finally the noise power indicated by the superhet system has to be determined.

If a carrier is simultaneously amplitude, and frequency, modulated by a sinusoidal signal, the a.m. upper (lower) sideband component can be written

$$[e_{u(1)}]_{\mathrm{am}} = \tfrac{1}{2}\hat{E}_o M \cos[\{\Omega(\pm)\omega\}\ t(\pm)\alpha]$$

where $\hat{E}_o$ is the peak amplitude of the carrier,

M is the a.m. modulation index

Ω, ω are the angular frequencies of the carrier, and the modulating signal, respectively,

and α is an arbitrary phase angle.

The f.m. upper (lower) sideband component, for the case of a very small f.m. index ($m \ll 1$), is:

$$[e_{u(1)}]_{\mathrm{fm}} = \tfrac{1}{2}\hat{E}_o m \cos[\{\Omega(\pm)\omega\}\ t(\pm)\alpha + \tfrac{1}{2}\pi]$$

At both the upper, and the lower, sideband frequencies there exist, therefore, two components which are in phase quadrature. The power resulting from two components in phase quadrature is equal to the sum of the powers of the individual components, as for the case of fluctuating quantities which are uncorrelated. Thus the superhet measuring system will indicate at the upper, as well as at the lower, sideband frequencies the power

$$N_s = k(\overline{A^2} + \overline{F^2} + \overline{B^2})$$

or

$$N_s = N_{\mathrm{am}} + N_{\mathrm{fm}} + N_{\mathrm{b}} \qquad \ldots\ldots(26)$$

The three noise powers $N_{d(\mathrm{am})}$, $N_{d(\mathrm{fm})}$, and N_s are those measured, respectively, by the a.m. and f.m. direct-detection systems and the superhet system. They are related to the three component noise powers by eqns. (24), (25) and (26), but unfortunately these equations are not independent since adding eqns. (24) and (25) gives

$$N_{d(\mathrm{am})} + N_{d(\mathrm{fm})} = 4N_s$$

Thus a direct calculation of the noise components from the results of the three types of measurement is not possible, unless one of them is known or can be ignored. For example at low sideband frequencies the background noise contribution is very small whereas at high sideband frequencies the a.m. noise is negligible. It is sometimes possible to estimate the background noise from measurements taken under pre-oscillation conditions.[29] Fortunately a separation of components is not always necessary, and the noise contribution of an oscillator to a given system can be measured by methods such as those described here.

7. Acknowledgments

The work described was carried out in the Department of Electronics, University of Southampton, and the authors would like to thank Professor E. E. Zepler for his interest and encouragement. Grateful acknowledgment is also made to Mr. N. W. W. Smith of the Mullard Research Laboratories for many useful discussions.

8. References

1. D. Middleton, W. M. Gottschalk and J. B. Wiesner, "Noise in cw magnetrons", *J. Appl. Phys.*, **24**, p. 1065, 1953.
2. D. Middleton, "Theory of phenomenological models and measurements of the fluctuating output of cw magnetrons", *Trans. Inst. Radio Engrs (Electron Devices)*, **ED-1**, p. 56, 1954.
3. J. R. Pierce, "The general sources of noise in vacuum tubes", *Trans. I.R.E.*, **ED-1**, p. 135, 1954.

4. L. D. Smullin and H. A. Haus, "Noise in Electron Devices", (John Wiley, New York, 1959).

5. J. B. Wiesner, "Source noise and its effect upon electronic systems", *Trans. I.R.E.*, **ED-1**, p. 3, 1954.

6. W. W. McLeod, Jr., "Microwave oscillator requirements for cw radar", *Trans. I.R.E.*, **ED-1**, p. 11, 1954.

7. A. van der Ziel, "Noise", (Chapman and Hall, London, 1955).

8. S. N. van Voorhis (Ed.), "Microwave Receivers" (McGraw-Hill, New York, 1948).

9. J. B. H. Kuper and M. C. Waltz, Work summarized in D. R. Hamilton, J. K. Knipp, and J. B. H. Kuper, "Klystrons and Microwave Triodes", chap. 17 (McGraw-Hill, New York, 1948).

10. G. C. Dalman and E. Ortiz, "Measurements of microwave local-oscillator noise", *Proc. Nat. Elec. Conf.*, **9**, p. 833, 1953.

11 R. Müller, "Noise measurements of microwave local oscillators", *Trans. I.R.E.*, **ED-1**, p. 42, 1954.

12. W. M. Gottschalk, "Direct detection measurements of noise in cw magnetrons", *Trans. I.R.E.*, **ED-1**, p. 91, 1954.

13. J. B. Cicchetti and J. Munushian, "Noise characteristics of a backward-wave oscillator", *I.R.E. National Conv. Rec.*, Part 3, p. 84, 1958.

14. D. E. T. F. Ashby and R. B. Dyott, "Measuring modulation noise from a high-power cw klystron amplifier", *Proc. Instn Elect. Engrs*, **105B**, p. 879, 1958.

15. W. P. N. Court, "A tunable filter for use in the measurement of excess noise from local oscillators", *Electronic Engng*, **30**, p. 208, 1958.

16. N. W. W. Smith, "Noise in backward-wave oscillators", *Proc. Instn Elect. Engrs*, **105B**, p. 800, 1958.

17. R. V. Pound, "Microwave Mixers" (McGraw-Hill, New York, 1948).

18. H. C. Torrey and C. A. Whitmer, "Crystal Rectifiers" (McGraw-Hill, New York, 1948).

19. M. E. Sprinks, G. T. G. Robinson, and B. G. Bosch, "The frequency dependence of noise temperature ratio in microwave mixer crystals", *Brit. J. Appl. Phys.*, **8**, p. 275, 1957.

20. B. G. Bosch and W. A. Gambling, unpublished work.

21. G. Convert and O. Doehler, "The signal to noise ratio in the M-carcinotron", *Trans. I.R.E.*, **ED-1**, p. 184, 1954.

22. R. A. La Plante, "Development of a low-noise X-band cw klystron power oscillator", *Trans. I.R.E.*, **ED-1**, p. 99, 1954.

23. G. C. Dalman and A. S. Rhoads, "Microwave oscillator noise spectrum measurements", *Trans. I.R.E.*, **ED-1**, p. 51, 1954.

24. R. Adler, "A study of locking phenomena in oscillators", *Proc. Inst. Radio Engrs*, **34**, p. 351, 1946.

25. R. D. Huntoon and A. Weiss, "Synchronization of oscillators", *Proc. Inst. Radio Engrs*, **35**, p. 1415, 1947.

26. C. S. Aitchison, "Frequency synchronization of an X-band reflex klystron", *Proc. Instn Elect. Engrs*, **105B**, p. 944, 1958.

27. J. C. R. Licklider and E. Dzendolet, "Oscillographic scatter plots illustrating various degrees of correlation", *Science*, **107**, p. 121, 1948.

28. G. R. Sugar, "Estimation of correlation coefficients from scatter diagrams", *J. Appl. Phys.*, **25**, p. 354, 1954.

29. B. G. Bosch and W. A. Gambling, "Noise measurements on low-power X-band oscillators", presented at International Congress on Microwave Tubes, Munich, 1960.

30. C. Colani, "A simple microwave discriminator", *Frequenz*, **10**, p. 25, 1956.

Manuscript received by the Institution on 7th November 1960. (*Paper No. 643.*)

Careers Guidance in Radio and Electronic Engineering

The Institution has for many years given advice to Youth Employment Officers, careers masters and school boys on the various methods of entry into the profession. The General Secretary and the Education Officer have given lectures and talks to school leavers and school science societies on this subject, as well as representing the Institution at careers evenings which are becoming an annual event at many schools. Quite apart from this, Local Sections take an active part in careers guidance and provide material for exhibitions, as well as attending careers evenings.

Other professional bodies and industry generally also provide help in their own fields of activity. However, it was to draw attention to careers prospects in engineering and applied science, and to the technical training facilities available in Colleges of Technology and industry, that all these organizations were brought together to produce a concentrated careers effort under the title of "Commonwealth Technical Training Week".

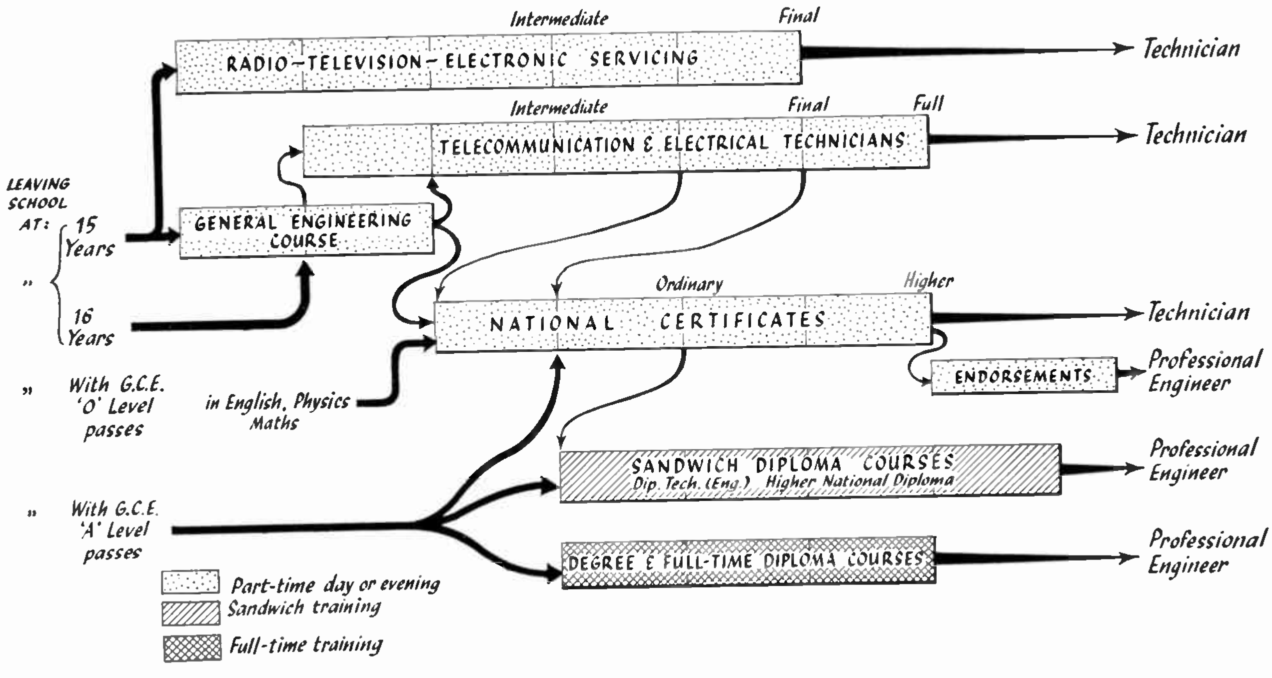
This was held from 29th May to 3rd June in the United Kingdom and a little earlier in other parts of the Commonwealth. The main aim of the Week's activities was to highlight and emphasize the importance of technical training in this technological age.

This venture was conceived by H.R.H. the Duke of Edinburgh, K.G., who took a very active part in the organization and in the preparations for this event. His Royal Highness attended several church services and official ceremonies in different parts of the country in connection with the Week. The efforts in various parts of the country and Commonwealth were co-ordinated by the City and Guilds of London Institute.

In the main the Commonwealth Technical Training Week took the form of exhibitions conducted by Colleges of Technology and technical colleges, who also held open days for parents and schoolchildren. In addition, this effort was supported by business organizations and civic authorities who organized central exhibitions in many of the larger towns.

The Commonwealth Technical Training Week was given wide publicity in the press and on radio and television, with the result that almost all local organizers were able to report that their activities were well supported.

For these displays the Institution provided copies of a booklet entitled "Careers in Radio or Electronic Engineering". This was supported by a wall chart showing the main career routes and methods of training available. A reproduction of the wall chart is given below. This material can be supplied to any school or college, and copies will be sent on request to members who are interested in careers guidance.



Career Routes in Radio and Electronic Engineering.

U.D.C. 512.83 : 621.375

A Matrix Representation of Linear Amplifiers

By

K. G. NICHOLS, M.Sc.,
(*Associate Member*)†

Summary: The transfer matrices of linear bilateral and unilateral networks are distinguished by the vanishing of the determinant of the matrix, in the case of the latter, and the non-existence of the matrix of the reversed four-pole. Illustrative examples are given. The voltage gain, current gain, input impedance and output impedance of a four-pole are deduced in terms of the transfer matrix elements. Expressions are developed for the transfer matrices of the four-poles resulting from combining pairs of four-poles, which may be either bilateral or unilateral, in parallel – parallel, in series – parallel and in series – series connection. These expressions are used to obtain the transfer matrices of some passive and active networks, including feedback amplifiers, to illustrate the possibilities of the technique.

1. Introduction

The matrix theory of linear passive four-poles, and their combination in parallel–parallel, series–parallel, parallel–series and series–series connections has been considered by many authors.[1,2] The extensions of the theory to active networks is usually developed in terms of the admittance matrix [3,4,6] or the transfer matrix [5,7] although in the latter case it has been usual to convert the transfer matrix into an admittance matrix when considering four-poles connected in the parallel–parallel arrangement.[5,6,7]

In the present paper the matrix theory of passive and active four-poles is developed exclusively in terms of the transfer matrices of the four-poles, the theory being valid for both unilateral and bilateral networks. Since the determinant of the transfer matrix of a unilateral four-pole is zero, the transfer matrix of the reversed four-pole does not exist. In a general theory it is therefore necessary to indicate a forward direction on each four-pole, in order to accommodate unilateral units. Distinction must be made between coincident and opposed forward directions when pairs of four-poles are combined in any of the usual four connections. Expressions are developed for the transfer matrices of pairs of four-poles combined in this way and are used to find the transfer matrices of illustrative examples.

2. Transfer Matrix of a Linear Four-Pole

With reference to Fig. 1 in which the arrow shows the input to output direction, if the input and output voltages and currents are related by,

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = \begin{pmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \qquad \ldots\ldots(1)$$

† Letchworth College of Technology, Hertfordshire.

where $\begin{pmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{pmatrix} = A$ and is independent of V_2 and i_2 ......(2)

then A is called the transfer matrix of the four-pole.

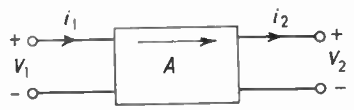


Fig. 1. A four-pole.

The actual network between the input and output terminals of the four-pole may consist of any arrangement of linear devices for which an equation of the form of eqn. (1) is valid. If the output of the four-pole is open-circuited then

$$i_2 = 0$$

and $$a_{11} = V_1/V_2; \qquad a_{21} = i_1/V_2 \qquad \ldots\ldots(3)$$

while if the output is short-circuited then

$$V_2 = 0$$

and $$a_{12} = V_1/i_2; \qquad a_{22} = i_1/i_2 \qquad \ldots\ldots(4)$$

The following terminology is useful:

$\dfrac{1}{a_{11}} = \dfrac{V_2}{V_1}$, output open-circuited, will be called the "voltage amplification factor";

$\dfrac{1}{a_{21}} = \dfrac{V_2}{i_1}$, output open-circuited, will be called the "mutual resistance factor";

$\dfrac{1}{a_{12}} = \dfrac{i_2}{V_1}$, output short circuited, will be called the "mutual conductance factor"; and

$\dfrac{1}{a_{22}} = \dfrac{i_2}{i_1}$, output short circuited, will be called the "current amplification factor"; of the four-pole.

$1/a_{11}$ is analogous to $-\mu$ for a grounded cathode triode amplifier and $1/a_{12}$ is analogous to $-g_m$ for the same amplifier, μ and g_m being the amplification factor and mutual conductance respectively of the triode.

3. The Transfer Matrix of the Reversed Four-Pole

The possibility of relating the output voltage and current linearly in terms of the input voltage and current depends on the existence of a unique inverse matrix A^{-1} of the transfer matrix A. Thus from eqn. (1),

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = A \begin{pmatrix} V_2 \\ i_2 \end{pmatrix}$$

whence
$$A^{-1}\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = A^{-1}A \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \qquad \ldots\ldots(5)$$

and so
$$\begin{pmatrix} V_2 \\ i_2 \end{pmatrix} = A^{-1} \begin{pmatrix} V_1 \\ i_1 \end{pmatrix} \qquad \ldots\ldots(6)$$

Since $\quad A^{-1}A = I\,(2\times 2 \text{ unit matrix}) \quad \ldots(7)$

then $\quad |A^{-1}A| = 1$

and so $\quad |A^{-1}||A| = 1$

Consequently the existence of an inverse of A necessarily implies that $|A| = a \neq 0$. That this condition is also sufficient for the existence of an inverse of A, follows since

$$\frac{1}{a}\begin{pmatrix} a_{22}, & -a_{12} \\ -a_{21}, & a_{11} \end{pmatrix}$$

is an inverse matrix of A as can be seen by direct multiplication with A. Further this is a unique inverse, for if X and Y are both inverses of A then,

$$XA = I \quad \text{and} \quad YA = I$$

On subtracting,

$$(X-Y)A = 0$$

which implies $\quad X = Y$

that is the inverse

$$A^{-1} = \frac{1}{a}\begin{pmatrix} a_{22}, & -a_{12} \\ -a_{21}, & a_{11} \end{pmatrix} \qquad \ldots\ldots(8)$$

is unique.

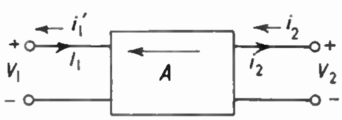


Fig. 2. The reversed four-pole.

In order to complete the formal interchange of the input and output terminals of the four-pole, it is necessary to reverse the input and output current senses as shown in Fig. 2. Equation (6) then becomes

$$\begin{pmatrix} V_2 \\ -i_2' \end{pmatrix} = \frac{1}{a}\begin{pmatrix} a_{22} & , -a_{12} \\ -a_{21} & , \quad a_{11} \end{pmatrix}\begin{pmatrix} V_1 \\ -i_1' \end{pmatrix} \qquad \ldots\ldots(9)$$

or
$$\begin{pmatrix} V_2 \\ i_2' \end{pmatrix} = \frac{1}{a}\begin{pmatrix} a_{22}, & a_{12} \\ a_{21}, & a_{11} \end{pmatrix}\begin{pmatrix} V_1 \\ i_1' \end{pmatrix} \qquad \ldots\ldots(10)$$

The matrix $\frac{1}{a}\begin{pmatrix} a_{22}, & a_{12} \\ a_{21}, & a_{11} \end{pmatrix}$, whenever it exists, is called the transfer matrix of the reversed four-pole, and it follows from the foregoing that the reversed four-pole is a linear four-pole, in the sense of the definition of eqn. (1), if and only if

$$a = |A| \neq 0 \qquad \ldots\ldots(11)$$

A four-pole will be called *unilateral* if the transfer matrix of the reverse four-pole does not exist, i.e. if $a = |A| = 0$, and will be called bilateral if its reverse four-pole transfer matrix exists, that is if $a = |A| \neq 0$.

In practical circuitry there is always some back coupling between output and input of a four-pole due to stray capacitances and stray mutual inductances; this implies that the transverse matrix of the reverse four-pole exists and so the device is bilateral. However, in many instances the degree of this back coupling is very small and the determinant of the transfer matrix approaches zero; in such cases it is usually convenient to put the determinant equal to zero and to consider the four-pole as unilateral.

It is worthy of note at this stage that the transfer matrix of a linear "passive" four-pole, i.e. one involving only passive elements, has

$$a = |A| = 1 \qquad \ldots\ldots(12)$$

The verification of this result is rather involved,[1] except in the case of a symmetrical passive network when the matrix of the four-pole and that of the reversed four-pole are identical, thus giving the result immediately.

The more general result is considered in Appendix 7.

4. Series Cascaded Four-Poles

Four-poles may be series cascaded thereby producing more complex four-poles; the transfer matrix of the resultant four-pole can be deduced from the transfer matrices of the cascaded four-poles as in Fig. 3. Here

$$\begin{pmatrix} V_2 \\ i_2 \end{pmatrix} = B \begin{pmatrix} V_3 \\ i_3 \end{pmatrix} \qquad \ldots\ldots(13)$$

and
$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = A \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \qquad \ldots\ldots(14)$$

Equations (13) and (14) then give

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = AB \begin{pmatrix} V_3 \\ i_3 \end{pmatrix} \qquad \ldots\ldots(15)$$

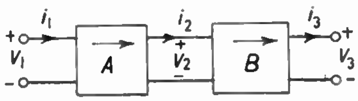


Fig. 3. Series cascaded four-poles.

where

$$AB = \begin{pmatrix} a_{11} , & a_{12} \\ a_{21} , & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} , & b_{12} \\ b_{21} , & b_{22} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} , & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} , & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix} \quad \ldots(16)$$

is the transfer matrix of the resultant four-pole.

5. The Transfer Matrices of some Simple Four-Poles

5.1. The transfer matrix of the series impedance shown in Fig. 4 is readily deduced as follows:

$$V_1 = V_2 + Zi_2$$

and

$$i_1 = 0.V_2 + i_2$$

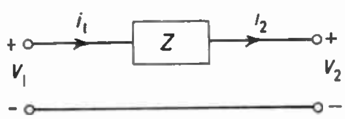


Fig. 4. Series impedance four-pole.

This gives the transfer matrix as

$$A = \begin{pmatrix} 1 , & Z \\ 0 , & 1 \end{pmatrix} \quad \ldots\ldots(17)$$

For this matrix $a = |A| = 1$, as for any passive four-pole, and the reversed four-pole has the same transfer matrix A.

5.2. The transfer matrix of the shunt admittance shown in Fig. 5 follows in a similar fashion.

$$V_1 = V_2 + 0.i_2$$

and

$$i_1 = YV_2 + i_2$$

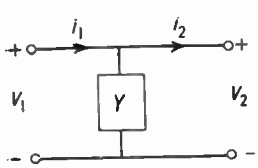


Fig. 5. Shunt admittance four-pole.

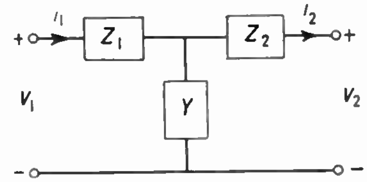
This gives

$$A = \begin{pmatrix} 1 , & 0 \\ Y , & 1 \end{pmatrix} \quad \ldots\ldots(18)$$

Again $a = |A| = 1$, and the reversed four-pole has the same transfer matrix.

5.3. The transfer matrix of a more complicated passive four-pole can be deduced by cascading elements such as those in 5.1 and 5.2. In this way the transfer matrix of the T-network shown in Fig. 6 can be deduced.

Fig. 6. T-network.



By repeated application of eqn. (16) the transfer matrix of this network is given by,

$$\begin{pmatrix} 1 , & Z_1 \\ 0 , & 1 \end{pmatrix} \begin{pmatrix} 1 , & 0 \\ Y , & 1 \end{pmatrix} \begin{pmatrix} 1 , & Z_2 \\ 0 , & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1+Z_1Y , & Z_1 \\ Y , & 1 \end{pmatrix} \begin{pmatrix} 1 , & Z_2 \\ 0 , & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1+Z_1Y , & Z_1+Z_2+YZ_1Z_2 \\ Y , & 1+Z_2Y \end{pmatrix} \quad \ldots(19)$$

5.4. The transfer matrices of simple active elements can often be deduced very easily. Thus those of the voltage and current generators shown in Figs. 7 and 8 respectively are computed below.

Fig. 7. An active four-pole. Output voltage proportional to input voltage.

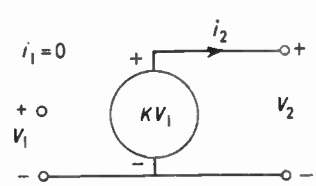
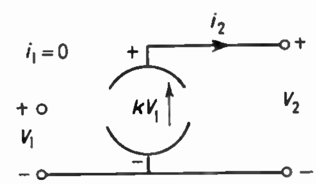


Fig. 8. An active four-pole. Output current proportional to input voltage.



For the arrangement of Fig. 7

$$V_1 = \frac{1}{K} V_2 + 0.i_2$$

and

$$i_1 = 0.V_2 + 0.i_2$$

The transfer matrix is thus

$$A = \begin{pmatrix} \frac{1}{K} , & 0 \\ 0 , & 0 \end{pmatrix} \quad \ldots\ldots(20)$$

Since $a = |A| = 0$, the transfer matrix of the reversed network does not exist, as is of course immediately obvious from the figure since i_2 is independent of both V_1 and i_1. This four-pole is therefore unilateral.

For the arrangement of Fig. 8

$$V_1 = 0.V_2 + \frac{1}{k}.i_2$$

and $$i_1 = 0.V_2 + 0.i_2$$

The transfer matrix is thus

$$A = \begin{pmatrix} 0 & , & \frac{1}{k} \\ 0 & , & 0 \end{pmatrix} \qquad \ldots\ldots(21)$$

Again this matrix has no inverse, and is unilateral.

5.5. The transfer matrix of the equivalent circuit of a linear class A1 grounded-cathode triode amplifier shown in Fig. 9 is readily obtained from the results obtained earlier.

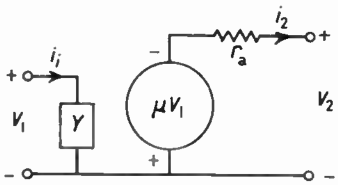


Fig. 9. Equivalent circuit of a common-cathode triode valve.

The transfer matrix of this four-pole is given by

$$A = \begin{pmatrix} 1 & , & 0 \\ Y & , & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\mu} & , & 0 \\ 0 & , & 0 \end{pmatrix} \begin{pmatrix} 1 & , & r_a \\ 0 & , & 1 \end{pmatrix} \ldots\ldots(22)$$

$$A = -\begin{pmatrix} \frac{1}{\mu} & , & \frac{1}{g_m} \\ \frac{Y}{\mu} & , & \frac{Y}{g_m} \end{pmatrix} \quad \text{where } \mu = r_a g_m \qquad \ldots\ldots(23)$$

The determinant of the transfer matrix vanishes, as is seen from eqn. (23), and so the transfer matrix has no inverse. The vanishing of the determinant can also be seen from eqn. (22) for

$$a = |A| = \begin{vmatrix} 1 & , & 0 \\ Y & , & 1 \end{vmatrix} \begin{vmatrix} -\frac{1}{\mu} & , & 0 \\ 0 & , & 0 \end{vmatrix} \begin{vmatrix} 1 & , & r_a \\ 0 & , & 1 \end{vmatrix}$$

$$= 1.0.1 = 0$$

thus showing that if any four-pole in a series cascaded set is unilateral then the resulting four-pole is unilateral.

5.6. As another example, the transfer matrix of the small-signal equivalent circuit for grounded base configuration of a transistor will be considered. The equivalent circuit in terms of the hybrid parameters (Fig. 10) will be used although any other form of equivalent circuit would serve equally well.

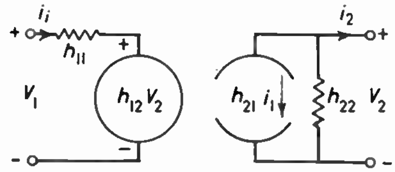


Fig. 10. Hybrid parameter equivalent circuit of a transistor. h_{11} ohms; h_{12} dimensionless; h_{21} dimensionless; h_{22} ohms.

From inspection of Fig. 10 it follows that

$$V_1 = h_{11}\, i_1 + h_{12}\, V_2 \qquad \ldots\ldots(24)$$

and $$-i_2 = h_{21}\, i_1 + h_{22}\, V_2 \qquad \ldots\ldots(25)$$

Elimination of i_1 gives,

$$h_{21}\, V_1 = (h_{12}\, h_{21} - h_{11}\, h_{22}) V_2 - h_{11}\, i_2$$

that is, $$V_1 = -\frac{h}{h_{21}} V_2 - \frac{h_{11}}{h_{21}} i_2 \qquad \ldots\ldots(26)$$

since $h_{21} \neq 0$ for any practical transistor, and where

$$h = |h| = h_{11}\, h_{22} - h_{12}\, h_{21}$$

Equation (25) may be written

$$i_1 = -\frac{h_{22}}{h_{21}} V_2 - \frac{1}{h_{21}} i_2 \qquad \ldots\ldots(27)$$

The transfer matrix of the four-pole is given by eqns. (26) and (27) as

$$A = -\frac{1}{h_{21}} \begin{pmatrix} h & , & h_{11} \\ h_{22} & , & 1 \end{pmatrix} \qquad \ldots\ldots(28)$$

The determinant of A is

$$\frac{1}{h_{21}^2}(h - h_{11}\, h_{22}) = \frac{h_{12}}{h_{21}} \qquad \ldots\ldots(29)$$

Since this is not in general zero, an inverse matrix exists, the transfer matrix of the reverse four-pole can be found, and so the four-pole is bilateral. In some transistors h_{12}/h_{21} is very small and the device is almost unilateral.

6. The Gain of a Linear Four-Pole

Figure 11 shows a four-pole with a load connected to its output terminals.

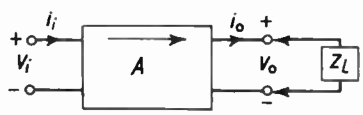


Fig. 11. A four-pole terminated with a load.

The load current and voltage are related by

$$V_o = Z_L . i_o \qquad \ldots\ldots(30)$$

Now, $$\begin{pmatrix} V_i \\ i_i \end{pmatrix} = \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix} \begin{pmatrix} V_o \\ i_o \end{pmatrix} \qquad \ldots\ldots(31)$$

and so, $$V_i = a_{11} V_o + a_{12} i_o$$
$$= a_{11} V_o + a_{12} V_o/Z_L$$

that is, the voltage gain of the four pole

$$= \frac{V_o}{V_i} = \frac{1}{a_{11} + a_{12}/Z_L} \quad \text{......(32)}$$

Similarly, the trans-resistance gain

$$= \frac{V_o}{i_i} = \frac{1}{a_{21} + a_{22}/Z_L} \quad \text{......(33)}$$

the trans-conductance gain

$$= \frac{i_o}{V_i} = \frac{1}{a_{11} Z_L + a_{12}} \quad \text{......(34)}$$

and the current gain $$= \frac{i_o}{i_i} = \frac{1}{a_{21} Z_L + a_{22}} \quad \text{......(35)}$$

Obviously eqn. (34) can be deduced from eqn. (32) by division by Z_L, and (35) can be deduced from (33) by division by Z_L.

To illustrate these results the matrix derived in Section 5.5 for the grounded cathode triode, namely

$$A = -\begin{pmatrix} \frac{1}{\mu} & , & \frac{1}{g_m} \\ \frac{Y}{\mu} & , & \frac{Y}{g_m} \end{pmatrix} \quad \text{......(23)}$$

will be used.

From this matrix the voltage gain of the grounded cathode triode amplifier is given by eqn. (32) as

$$\frac{1}{a_{11} + a_{12}/Z_L} = -\frac{1}{\frac{1}{\mu} + \frac{1}{g_m Z_L}} = -\frac{\mu Z_L}{Z_L + r_a} \quad \text{...(36)}$$

A second illustration is the matrix of the grounded-base transistor amplifier of Section 5.6. This matrix is

$$A = -\frac{1}{h_{21}} \begin{pmatrix} h & , & h_{11} \\ h_{22} & , & 1 \end{pmatrix} \quad \text{......(28)}$$

The voltage and current gains are then, respectively,

$$\frac{1}{a_{11} + a_{12}/Z_L} = -\frac{h_{21}}{h + h_{11}/Z_L} \quad \text{......(37)}$$

and $$\frac{1}{a_{21} Z_L + a_{22}} = -\frac{h_{21}}{h_{22} Z_L + 1} \quad \text{......(38)}$$

The gains defined by eqns. (32) to (35) are best known as the "device" gains and, while of use, are not generally as important as the "stage" gains. To compute the "stage" gains it is necessary to consider the internal impedance of the voltage source, or the internal admittance of the current source, feeding the input terminals of the four-pole. Figure 12 shows a four-pole fed from voltage generators of internal source impedance Z_s, while Fig. 13 is a four-pole fed from current generator of internal source admittance Y_s.

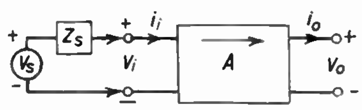


Fig. 12. A four-pole driven from a voltage source.

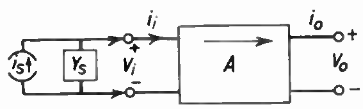


Fig. 13. A four-pole driven from a current source.

In either case the effect of the source impedance or admittance can be allowed for by the rule of series cascading four-poles (section 4). Thus for the series source impedance Z_s, the transfer matrix of the cascaded four-poles is

$$\begin{pmatrix} 1 & , & Z_s \\ 0 & , & 1 \end{pmatrix} \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11} + Z_s . a_{21} & , & a_{12} + Z_s . a_{22} \\ a_{21} & , & a_{22} \end{pmatrix} \quad \text{......(39)}$$

While for the shunt source admittance Y_s, the transfer matrix of the cascaded four-poles is

$$\begin{pmatrix} 1 & , & 0 \\ Y_s & , & 0 \end{pmatrix} \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11} & , & a_{12} \\ Y_s . a_{11} + a_{21} & , & Y_s . a_{12} + a_{22} \end{pmatrix} \quad \text{......(40)}$$

The "stage" gains can now be evaluated using the transfer matrices given by eqns. (39) and (40) in the "device" gain eqns. (32) to (35).

7. The Input Impedance of a Four-Pole

Referring to Fig. 11, in which the output of the four-pole is terminated by a load impedance Z_L, the input impedance is given by, from eqn. (31), as

$$Z_{in} = \frac{V_i}{i_i} = \frac{a_{11} V_o + a_{12} i_o}{a_{21} V_o + a_{22} i_o}$$

with $$V_o = Z_L . i_o \quad \text{......(30)}$$

that is, $$Z_{in} = \frac{a_{11} . Z_L + a_{12}}{a_{21} Z_L + a_{22}} \quad \text{......(41)}$$

7.1. As an example, the grounded cathode amplifier has transfer matrix

$$A = -\begin{pmatrix} \dfrac{1}{\mu} & , & \dfrac{1}{g_m} \\ \dfrac{Y}{\mu} & , & \dfrac{Y}{g_m} \end{pmatrix} \qquad \ldots\ldots(23)$$

is considered. For this,

$$Z_{\text{in}} = \frac{1/\mu . Z_L + 1/g_m}{Y/\mu . Z_L + Y/g_m} = \frac{1}{Y}$$

as might be expected from more elementary considerations.

7.2. Again, as a second example, the grounded base transistor amplifier considered earlier, has transfer matrix

$$A = -\frac{1}{h_{21}}\begin{pmatrix} h & , & h_{11} \\ h_{22} & , & 1 \end{pmatrix} \qquad \ldots\ldots(28)$$

giving

$$Z_{\text{in}} = \frac{h . Z_L + h_{11}}{h_{22} . Z_L + 1}$$

This may also be written as

$$h_{11} - \frac{h_{12} h_{21} Z_L}{h_{22} Z_L + 1}$$

which is a better known form.

8. The Output Impedance of a Four-Pole

The output impedance of a four-pole can be determined by evaluating the quotient $V_o/-i_o$ after short-circuiting the input voltage source or open-circuiting the input current source. Figures 12 and 13 then become identical and the output impedance will be given by an expression similar to eqn. (41), in terms of the transfer matrix of the reversed four-pole, if the four-pole is bilateral. Assuming for the present that this latter condition is true, the transfer matrix of the reversed four-pole exists and is given by eqn. (10) as,

$$B = \begin{pmatrix} b_{11} & , & b_{12} \\ b_{21} & , & b_{22} \end{pmatrix} = \frac{1}{a}\begin{pmatrix} a_{22} & , & a_{12} \\ a_{21} & , & a_{11} \end{pmatrix} \quad \ldots(42)$$

The output impedance of the four-pole is then, by analogy with eqn. (41),

$$Z_{\text{out}} = \frac{a_{22} Z_s + a_{12}}{a_{21} Z_s + a_{11}} \qquad \ldots\ldots(43)$$

8.1. The output impedance of the grounded base transistor amplifier, considered earlier, is

$$Z_{\text{out}} = \frac{Z_s + h_{11}}{h_{22} . Z_s + h}$$

on referring to eqn. (28), which may also be written as

$$Z_{\text{out}} = \frac{1}{h_{22} - \dfrac{h_{12} h_{21}}{Z_s + h_{11}}}$$

a better known form.

In fact eqn. (43) is valid even for a unilateral four-pole, that is one for which the transfer matrix of the reverse four-pole does not exist. This can be shown as follows:

$$\begin{pmatrix} V_i \\ i_i \end{pmatrix} = \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix}\begin{pmatrix} V_o \\ i_o \end{pmatrix} \qquad \ldots\ldots(31)$$

Now $\quad Z_s = V_i/-i_s$

for the input voltage generator shorted or input current-generator open circuited, and

$$Z_{\text{out}} = V_o/-i_o$$

Hence, from eqn. (31)

$$-Z_s = \frac{-a_{11} . Z_{\text{out}} + a_{12}}{-a_{21} . Z_{\text{out}} + a_{22}}$$

This gives

$$-a_{21} . Z_s . Z_{\text{out}} + a_{22} . Z_s = a_{11} . Z_{\text{out}} - a_{12}$$

that is

$$Z_{\text{out}} = \frac{a_{22} . Z_s + a_{12}}{a_{21} . Z_s + a_{11}} \qquad \ldots\ldots(43)$$

This derivation is independent of the non-vanishing or otherwise of $a = |A|$, and so is valid for both bilateral and unilateral four-poles.

8.2. The grounded cathode triode amplifier with transfer matrix

$$A = -\begin{pmatrix} \dfrac{1}{\mu} & , & \dfrac{1}{g_m} \\ \dfrac{Y}{\mu} & , & \dfrac{Y}{g_m} \end{pmatrix} \qquad \ldots\ldots(28)$$

can be used as an example of a unilateral four-pole. For this four-pole eqn. (43) gives the output impedance as

$$Z_{\text{out}} = \frac{YZ_s/g_m + 1/g_m}{YZ_s/\mu + 1/\mu}$$

$$= r_a\left(\frac{YZ_s + 1}{YZ_s + 1}\right) = r_a$$

as is known from more elementary considerations.

9. Four-Poles in Parallel–Parallel

In order to deal with certain feedback systems, and as an aid to building up transfer matrices of complicated four-poles, a rule for finding the transfer matrix of two four-poles in parallel is now required.

Referring to Fig. 14,

$$\begin{pmatrix} V_1 \\ i_{1A} \end{pmatrix} = \begin{pmatrix} a_{11} , & a_{12} \\ a_{21} , & a_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_{2A} \end{pmatrix} \qquad \ldots\ldots(44)$$

$$\begin{pmatrix} V_1 \\ i_{1B} \end{pmatrix} = \begin{pmatrix} b_{11} , & b_{12} \\ b_{21} , & b_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_{2B} \end{pmatrix} \qquad \ldots\ldots(45)$$

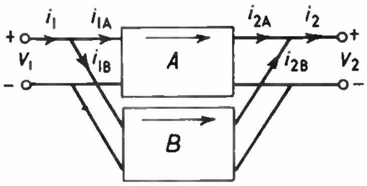


Fig. 14. Four-poles in parallel–parallel connection. (The same forward directions).

and from these equations it is shown in Appendix 1 that the transfer matrix of the complete four-pole is given as,

$$C = \frac{1}{a_{12}+b_{12}} \begin{pmatrix} a_{12}b_{11}+a_{11}b_{12} , & a_{12}b_{12} \\ (a_{12}+b_{12})(a_{21}+b_{21})-(a_{11}-b_{11})(a_{22}-b_{22}) , & a_{22}b_{12}+b_{22}a_{12} \end{pmatrix} \qquad \ldots\ldots(46)$$

provided $a_{12}+b_{12} \neq 0$.

9.1. As an example of the use of eqn. (46), the transfer matrix of the bridged-T network shown in Fig. 15 is deduced.

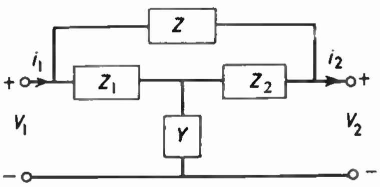


Fig. 15. Bridged-T-network.

The transfer matrix of the "T" is

$$\begin{pmatrix} 1 , & Z_1 \\ 0 , & 1 \end{pmatrix} \begin{pmatrix} 1 , & 0 \\ Y , & 1 \end{pmatrix} \begin{pmatrix} 1 , & Z_2 \\ 0 , & 1 \end{pmatrix} = \begin{pmatrix} 1+Z_1Y , & Z_1+Z_2+YZ_1Z_2 \\ Y , & 1+Z_2Y \end{pmatrix}$$

Using eqn. (46) the transfer matrix of the bridged-T becomes

$$\frac{1}{Z+Z_1+Z_2+YZ_1Z_2} \begin{pmatrix} Z+ZZ_1Y+Z_1+Z_2+YZ_1Z_2 , & ZZ_1+ZZ_2+YZZ_1Z_2 \\ (Z+Z_1+Z_2+YZ_1Z_2)Y-Z_1Z_2Y^2 , & Z+ZZ_2Y+Z_1+Z_2+YZ_1Z_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1+\dfrac{ZZ_1Y}{Z+Z_1+Z_2+YZ_1Z_2} , & Z-\dfrac{Z^2}{Z+Z_1+Z_2+YZ_1Z_2} \\ \dfrac{(Z+Z_1+Z_2)Y}{Z+Z_1+Z_2+YZ_1Z_2} , & 1+\dfrac{ZZ_2Y}{Z+Z_1+Z_2+YZ_1Z_2} \end{pmatrix}$$

9.2. As a further example of the use of eqn. (46), the operational amplifier shown in Fig. 16 is considered.

The unilateral amplifier has an unloaded gain or amplification factor $-G$, an output impedance Z_o, and an input admittance Y_i.

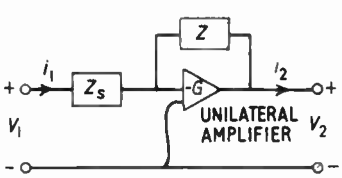


Fig. 16. An operational amplifier.

The transfer matrix of this amplifier is

$$\begin{pmatrix} 1 , & 0 \\ Y_i , & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{G} , & 0 \\ 0 , & 0 \end{pmatrix} \begin{pmatrix} 1 , & Z_o \\ 0 , & 1 \end{pmatrix} = -\frac{1}{G} \begin{pmatrix} 1 , & Z_o \\ Y_i , & Z_oY_i \end{pmatrix}$$

Then using eqn. (46) the transfer matrix of the amplifier shunted by Z is

$$-\frac{1}{GZ-Z_o}\begin{pmatrix} Z+Z_o & , & ZZ_o \\ Y_i\left(\dfrac{GZ-Z_o}{G}\right)+\dfrac{(G+1)(G+Y_iZ_o)}{G} & , & (1+Y_iZ)Z_o \end{pmatrix}$$

If this is series cascaded with the transfer matrix of Z_s the transfer matrix of the complete operational amplifier becomes

$$-\begin{pmatrix} \dfrac{Z+Z_o}{GZ-Z_o}+\dfrac{Z_sY_i}{G}+\dfrac{Z_s(G+1)(G+Y_iZ_o)}{G(GZ-Z_o)} & , & Z_o(Z+Z_s(1+Y_iZ)) \\ \dfrac{Y_i}{G}+\dfrac{(G+1)(G+Y_iZ_o)}{G(GZ-Z_o)} & , & \dfrac{Z_o(1+Y_iZ)}{GZ-Z_o} \end{pmatrix}$$

From this transfer matrix the voltage or current gain, for any given load impedance, can be readily deduced. The input and output impedances, for given load and source impedances respectively, are also readily obtainable from this transfer matrix.

In order to obtain from this matrix an elementary recognizable result, it will be assumed that Z_o is negligible compared with GZ, that Y_iZ_o is negligible compared with G, and that the term Z_sY_i/G is negligible compared with the other terms of a_{11}. The transfer matrix then reduces to

$$-\frac{1}{GZ}\begin{pmatrix} Z+Z_o+Z_s(G+1) & , & (Z+Z_s)Z_o(1+Y_iZ) \\ G+1 & , & Z_o(1+Y_iZ) \end{pmatrix}$$

For sake of example, the voltage gain of the unloaded operational amplifier is then given as

$$\frac{1}{a_{11}}=-\frac{GZ}{Z+Z_o+Z_s(G+1)}$$

If $(Z+Z_o)$ is negligible compared with $Z_s(G+1)$ in the denominator of this last expression this result reduces to

$$\frac{1}{a_{11}}=-\frac{Z}{Z_s\left(1+\dfrac{1}{G}\right)}\simeq-\frac{Z}{Z_s}$$

for large $|G|$ compared with unity.

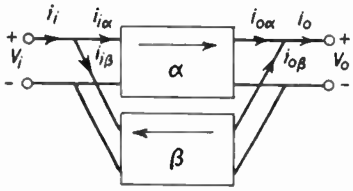


Fig. 17. Four-poles in parallel–parallel connection. (Opposed forward directions).

....

The method used in this latter example is clearly not restricted to unilateral amplifiers, and may be used for transistor operational amplifiers.

10. Shunt Feedback Applied in Parallel with the Input to the Amplifier

Figure 17 shows a feedback amplifier with a forward gain path represented by a four-pole with the transfer matrix α. The feedback path is represented by the four-pole with the transfer matrix β such that,

$$\begin{pmatrix} V_o \\ -i_{o\beta} \end{pmatrix}=\begin{pmatrix} \beta_{11} & , & \beta_{12} \\ \beta_{21} & , & \beta_{22} \end{pmatrix}\begin{pmatrix} V_i \\ -i_{i\beta} \end{pmatrix} \quad \ldots\ldots(47)$$

If the feedback path is bilateral, and in particular if it is passive, eqn. (47) may be written as

$$\begin{pmatrix} V_i \\ i_{i\beta} \end{pmatrix}=\begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix}\begin{pmatrix} V_o \\ i_{o\beta} \end{pmatrix} \quad \ldots\ldots(48)$$

where

$$\begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix}=\frac{1}{\beta}\begin{pmatrix} \beta_{22} & , & \beta_{12} \\ \beta_{21} & , & \beta_{11} \end{pmatrix} \quad \ldots\ldots(49)$$

is the transfer matrix of the reversed feedback four-pole.

Under this condition Fig. 17 shows two four-poles in parallel and the results of Section 9 apply, eqn. (46) then showing how to obtain the transfer matrix of the entire feedback amplifier. Indeed Section 9.2 concerned an example of a feedback amplifier of this kind.

When the feedback path is unilateral this technique does not apply because the reverse transfer matrix does not exist and it is necessary to work directly with eqn. (47). It is shown in Appendix 2 that the transfer matrix of the four-pole of Fig. 17 is given by,

$$C = \frac{1}{\beta_{12}+\alpha_{12}\beta}\begin{pmatrix} \alpha_{11}\beta_{12}+\alpha_{12}\beta_{22} & , & \alpha_{12}\beta_{12} \\ \alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-\alpha\beta-1 & , & \alpha_{12}\beta_{11}+\alpha_{22}\beta_{12} \end{pmatrix} \quad \ldots\ldots(50)$$

It is easily shown that this reduces to eqn. (46), a rather simpler matrix, if the feedback network is bilateral. For this reason eqn. (50) need only be used when the feedback network is unilateral. With this restriction $\beta = 0$ and transfer matrix of eqn. (50) reduces to

$$C = \frac{1}{\beta_{12}}\begin{pmatrix} \alpha_{11}\beta_{12}+\alpha_{12}\beta_{22} & , & \alpha_{12}\beta_{12} \\ \alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-1 & , & \alpha_{12}\beta_{11}+\alpha_{22}\beta_{12} \end{pmatrix} \quad \ldots\ldots(51)$$

11. The Combination of Four-Poles with Inputs in Series and Outputs in Parallel

To facilitate the analysis of other types of feedback amplifier it is desirable to be able to obtain the transfer matrix of a system of two four-poles with inputs in series and outputs in parallel. Such a system is shown in Fig. 18.

In this figure it will be observed that the current i_1 is assumed to flow out of the lower input terminal of the system. This assumption is certainly valid if the input to the system is in series cascade with another four-pole but is not necessarily valid in general.[1, 2] However since the input to the system will almost invariably be a voltage or current source, that is four-pole series cascaded with the four-pole of Fig. 18, the assumption is not unduly restrictive.

For the system of Fig. 18 the following equations are valid,

$$\begin{pmatrix} V_{1A} \\ i_1 \end{pmatrix} = \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix}\begin{pmatrix} V_2 \\ i_{2A} \end{pmatrix} \quad \ldots\ldots(52)$$

and

$$\begin{pmatrix} V_{1B} \\ -i_1 \end{pmatrix} = \begin{pmatrix} b_{11} & , & b_{12} \\ b_{21} & , & b_{22} \end{pmatrix}\begin{pmatrix} V_2 \\ i_{2B} \end{pmatrix} \quad \ldots\ldots(53)$$

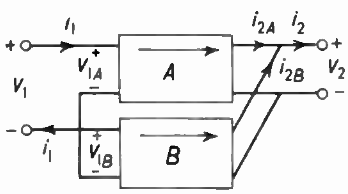


Fig. 18. Four-poles in series–parallel connection. (The same forward directions).

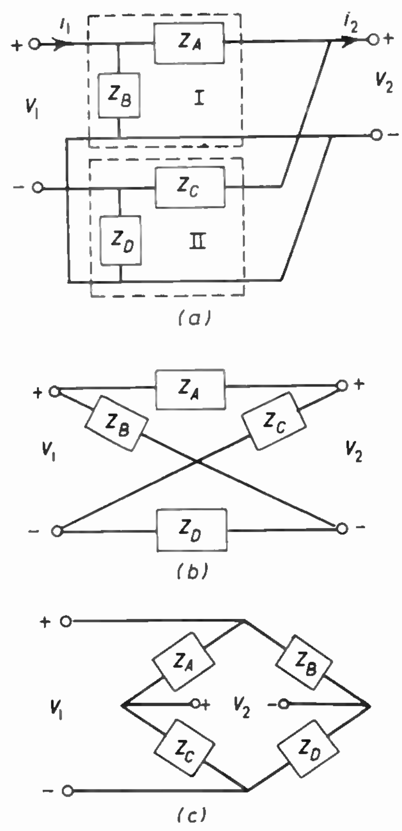


Fig. 19. (*a*) System of four-poles; (*b*) Rearrangement of (*a*) as an unbolted lattice (*c*) Rearrangement of (*b*) as a Wheatstone net.

11.1. The transfer matrix of the system of four-poles shown in Fig. 19(a) is readily obtained using eqn. (54).

From eqns. (52) and (53) it is shown in Appendix 3 that the transfer matrix of the complete system is

$$C = \frac{1}{b_{22}-a_{22}}\begin{pmatrix} (a_{12}+b_{12})(a_{21}+b_{21})-(a_{11}-b_{11})(a_{22}-b_{22}) & , & a_{12}b_{22}+b_{12}a_{22} \\ a_{21}b_{22}+b_{21}a_{22} & , & a_{22}b_{22} \end{pmatrix} \quad \ldots\ldots(54)$$

The transfer matrix of the four-pole I of Fig. 19(a) is

$$\begin{pmatrix} 1 & , & 0 \\ \dfrac{1}{Z_B} & , & 1 \end{pmatrix}\begin{pmatrix} 1 & , & Z_A \\ 0 & , & 1 \end{pmatrix}=\begin{pmatrix} 1 & , & Z_A \\ \dfrac{1}{Z_B} & , & 1+\dfrac{Z_A}{Z_B} \end{pmatrix}$$

and that of the four-pole II of the same figure is

$$\begin{pmatrix} 1 & , & Z_C \\ \dfrac{1}{Z_D} & , & 1+\dfrac{Z_C}{Z_D} \end{pmatrix}$$

Equation (54) then gives the transfer matrix of the complete network as,

$$\frac{1}{Z_B Z_C - Z_A Z_D}\begin{pmatrix} (Z_A+Z_C)(Z_B+Z_D) & , & Z_A Z_B Z_C + Z_B Z_C Z_D + Z_C Z_D Z_A + Z_D Z_A Z_B \\ Z_A+Z_B+Z_C+Z_D & , & (Z_A+Z_B)(Z_C+Z_D) \end{pmatrix}$$

Figure 19(a) can be rearranged to give the four-pole of Fig. 19(b), namely the unbolted lattice network, which can in turn be arranged to give the Wheatstone net of Fig. 19(c).

The transfer matrix of the network, used in conjunction with the gain eqns. (32) to (35), shows that all the gains, voltage, trans-resistance, trans-conductance or current, are zero, for any load impedance Z_L connected to the output, if

$$Z_B Z_C - Z_A Z_D = 0$$

This is of course the balance condition of the Wheatstone net. However the transfer matrix above gives much more information concerning the out of balance conditions of the network.

12. Shunt Feedback Applied in Series with the Input to the Amplifier

Figure 20 shows a feedback amplifier with a forward gain path represented by a four-pole with the transfer matrix α. The feedback path is represented by the four-pole with the transfer matrix β such that

$$\begin{pmatrix} V_o \\ -i_{o\beta} \end{pmatrix}=\begin{pmatrix} \beta_{11} & , & \beta_{12} \\ \beta_{21} & , & \beta_{22} \end{pmatrix}\begin{pmatrix} V_{i\beta} \\ i_i \end{pmatrix} \qquad \ldots\ldots(55)$$

If the feedback path is bilateral, and in particular if it is passive, eqn. (55) may be written as,

$$\begin{pmatrix} V_{i\beta} \\ -i_i \end{pmatrix}=\begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix}\begin{pmatrix} V_o \\ i_{o\beta} \end{pmatrix} \qquad \ldots\ldots(56)$$

where $$\begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix}=\frac{1}{\beta}\begin{pmatrix} \beta_{22} & , & \beta_{12} \\ \beta_{21} & , & \beta_{11} \end{pmatrix} \qquad \ldots\ldots(57)$$

is the transfer matrix of the reversed feedback four-pole.

Under this condition Fig. 20 shows two four-poles in the connection of Fig. 18 and the results of Section 11 are valid, eqn. (54) then giving the transfer matrix of the feedback amplifier.

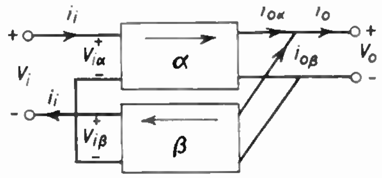


Fig. 20. Four-poles in series–parallel connection. (Opposed forward directions).

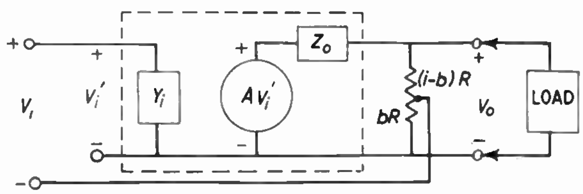


Fig. 21. Unilateral amplifier with voltage feedback in series with input.

When the feedback path is unilateral this method fails because the reverse transfer matrix does not exist and it is necessary to work directly with eqn. (55) to establish the transfer function of the entire amplifier. The details of this work are carried out in Appendix 4, where it is shown that the transfer matrix of the amplifier is

$$C=\frac{1}{\beta_{11}-\alpha_{22}\beta}\begin{pmatrix} \alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-\alpha\beta-1 & , & \alpha_{12}\beta_{11}+\alpha_{22}\beta_{12} \\ \alpha_{21}\beta_{11}+\alpha_{22}\beta_{21} & , & \alpha_{22}\beta_{11} \end{pmatrix} \qquad \ldots\ldots(58)$$

It is readily shown that this reduces to eqn. (54) when the feedback path is bilateral.

12.1. As an illustration of the use of eqn. (58) the unilateral feedback amplifier shown in Fig. 21 will be considered.

The unilateral amplifier has an input admittance Y_i, output impedance Z_o and unloaded voltage gain or amplification factor A. Thus the α matrix is

$$\begin{pmatrix} 1 & , & 0 \\ Y_i & , & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{A} & , & 0 \\ 0 & , & 0 \end{pmatrix}\begin{pmatrix} 1 & , & Z_o \\ 0 & , & 1 \end{pmatrix} = \frac{1}{A}\begin{pmatrix} 1 & , & Z_o \\ Y_i & , & Y_i Z_o \end{pmatrix}$$

The feedback voltage is obtained by taking a fraction b of the output voltage developed across a resistor R. The transfer matrix of the β feedback network is

$$\begin{pmatrix} 1 & , & (1-b)R \\ 0 & , & 1 \end{pmatrix}\begin{pmatrix} 1 & , & 0 \\ \frac{1}{bR} & , & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{b} & , & (1-b)R \\ \frac{1}{bR} & , & 1 \end{pmatrix}$$

Equation (58) then gives the transfer matrix of the amplifier as

$$\frac{1}{A - bY_i Z_o}\begin{pmatrix} 1 - bA + \frac{Z_o}{R} + bY_i(Z_o + (1-b)R) & , & Z_o(1 + bY_i(1-b)R) \\ Y_i\left(1 + \frac{Z_o}{R}\right) & , & Y_i Z_o \end{pmatrix}$$

If Y_i is taken as zero and R is very large compared with $|Z_o|$ this matrix reduces to

$$\begin{pmatrix} \frac{1-bA}{A} & , & \frac{Z_o}{A} \\ 0 & , & 0 \end{pmatrix}$$

The unloaded feedback amplifier gain is then

$$\frac{1}{a_{11}} = \frac{A}{1-bA}$$

and its output impedance is

$$\frac{a_{12}}{a_{11}} = \frac{Z_o}{1-bA}$$

12.2. The above example was used simply to illustrate the power of the method which is not restricted to unilateral α-path amplifiers.

13. The Combination of Four-Poles with Inputs and Outputs in Series

Another arrangement for combining four-poles is that shown in Fig. 22. For this system of four-poles

$$\begin{pmatrix} V_{1A} \\ i_1 \end{pmatrix} = \begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix}\begin{pmatrix} V_{2A} \\ i_2 \end{pmatrix} \quad \text{......(59)}$$

and

$$\begin{pmatrix} V_{1B} \\ -i_1 \end{pmatrix} = \begin{pmatrix} b_{11} & , & b_{12} \\ b_{21} & , & b_{22} \end{pmatrix}\begin{pmatrix} V_{2B} \\ -i_2 \end{pmatrix} \quad \text{......(60)}$$

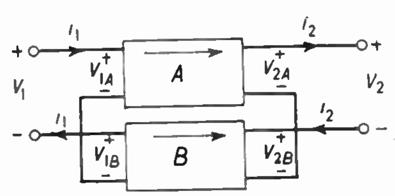


Fig. 22. Four-poles in series–series connection. (Same forward directions).

From these equations it is shown in Appendix 5 that the transfer matrix of the complete four-pole of Fig. 22 is

$$C = \frac{1}{a_{21}+b_{21}}\begin{pmatrix} a_{11}b_{21} + a_{21}b_{11} & , & (a_{12}+b_{12})(a_{21}+b_{21}) - (a_{11}-b_{11})(a_{22}-b_{22}) \\ a_{21}b_{21} & , & a_{22}b_{21} + a_{21}b_{22} \end{pmatrix} \quad \text{......(61)}$$

13.1. As an illustration of the use of eqn. (61) a grounded emitter transistor amplifier with undecoupled emitter resistance will be considered. The basic circuit of such an amplifier is shown in Fig. 23, in which

$$R_b = \frac{R_{b_1} . R_{b_2}}{R_{b_1} + R_{b_2}}$$

The hybrid parameters for the grounded emitter incremental equivalent circuit of the transistor will be taken as h_{11}, h_{12}, h_{21} and h_{22}, the prime or suffix e being omitted for convenience.

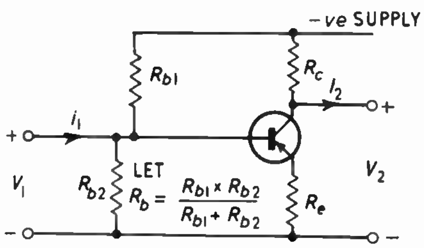


Fig. 23. Common emitter amplifier stage with undecoupled emitter resistor.

The transfer matrix of a grounded emitter transistor is (see Sect. 5.4)

$$A = -\frac{1}{h_{21}} \begin{pmatrix} h & , & h_{11} \\ h_{22} & , & 1 \end{pmatrix}$$

The emitter resistance R_e can be considered as a four-pole with transfer matrix

$$B = \begin{pmatrix} 1 & , & 0 \\ \frac{1}{R_e} & , & 1 \end{pmatrix}$$

which is combined with the transistor four-pole as in Fig. 22. The transfer matrix of the resultant four-pole is given by eqn. (61) as

$$\frac{1}{h_{22} R_e - h_{21}} \begin{pmatrix} h + h_{22} R_e & , & h_{11} + (h - h_{12} + h_{21} + 1) R_e \\ h_{22} & , & 1 + h_{22} R_e \end{pmatrix}$$

This four-pole is shunted at its input by a four-pole with transfer matrix

$$\begin{pmatrix} 1 & , & 0 \\ \frac{1}{R_b} & , & 1 \end{pmatrix}$$

and at its output by a four-pole with transfer matrix

$$\begin{pmatrix} 1 & , & 0 \\ \frac{1}{R_c} & , & 1 \end{pmatrix}$$

Series cascading these four-poles gives the transfer matrix of the whole amplifier stage as

$$\begin{pmatrix} 1 & , & 0 \\ \frac{1}{R_b} & , & 1 \end{pmatrix} \frac{1}{(h_{22} R_e - h_{21})} \cdot \begin{pmatrix} h + h_{22} R_e & , & h_{11} + (h - h_{12} + h_{21} + 1) R_e \\ h_{22} & , & 1 + h_{22} R_e \end{pmatrix} \begin{pmatrix} 1 & , & 0 \\ \frac{1}{R_c} & , & 1 \end{pmatrix}$$

$$= \frac{1}{h_{22} R_e - h_{21}} \begin{pmatrix} h + h_{22} R_e + \frac{h_{11}}{R_c} + (h - h_{12} + h_{21} + 1) \frac{R_e}{R_c} & , & h_{11} + (h - h_{12} + h_{21} + 1) R_e \\ \frac{h}{R_b} + h_{22} \left(\frac{R_e}{R_b} + \frac{R_e}{R_c} + 1 \right) + \frac{h_{11}}{R_b R_c} + (h - h_{12} + h_{21} + 1) \frac{R_e}{R_b R_c} + \frac{1}{R_c} & , & \frac{h_{11}}{R_b} + (h - h_{12} + h_{21} + 1) \frac{R_e}{R_b} + 1 + h_{22} R_e \end{pmatrix}$$

From this transfer matrix the voltage and current gains, and input and output impedances, for given load and source impedances, can be readily deduced. For example, the unloaded amplifier voltage gain is

$$\frac{1}{a_{11}} = -\frac{(h_{21} - h_{22} R_e) R_c}{(h + h_{22} R_e) R_c + h_{11} + (h - h_{12} + h_{21} + 1) R_e}$$

If $R_e = 0$, this reduces to

$$-\frac{h_{21}}{hR_c + h_{11}}$$

the usual gain expression.

14. Series Feedback Applied in Series with the Input to the Amplifier

Figure 24 shows a feedback amplifier with a forward gain path represented by a four-pole with the transfer matrix α. The feedback path is represented by the four-pole with the transfer matrix β such that

$$\begin{pmatrix} V_{o\beta} \\ i_o \end{pmatrix} = \begin{pmatrix} \beta_{11} & , & \beta_{12} \\ \beta_{21} & , & \beta_{22} \end{pmatrix} \begin{pmatrix} V_{i\beta} \\ i_i \end{pmatrix} \qquad \ldots\ldots(62)$$

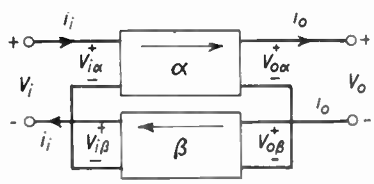


Fig. 24. Four-poles in series–series connection. (Opposed forward directions).

As in previous cases, if the feedback path is bilateral, eqn. (62) may be written as

$$\begin{pmatrix} V_{i\beta} \\ -i_i \end{pmatrix} = \begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix} \begin{pmatrix} V_{o\beta} \\ -i_o \end{pmatrix} \qquad \ldots\ldots(63)$$

where
$$\begin{pmatrix} \beta'_{11} & , & \beta'_{12} \\ \beta'_{21} & , & \beta'_{22} \end{pmatrix} = \frac{1}{\beta} \begin{pmatrix} \beta_{22} & , & \beta_{12} \\ \beta_{21} & , & \beta_{11} \end{pmatrix} \qquad \ldots\ldots(64)$$

is the transfer matrix of the reversed feedback four-pole.

Under this condition Fig. 24 shows two four-poles in the connection of Fig. 22 and the results of Sect. 13 are valid, eqn. (61) then gives the transfer matrix of the feedback amplifier.

Again, as in previous cases, when the feedback path is unilateral eqn. (62) must be used to deduce the transfer matrix of Fig. 24. This work is carried out in Appendix 6 where it is shown that the transfer matrix is given by

15. The Combination of Four-Poles with Inputs in Parallel and Outputs in Series and Series Feedback Applied in Shunt with the Input to the Amplifier

These cases appear to have little application in practice and so transfer matrices for such networks will not be derived here. The derivations of such transfer matrices are similar to those carried out in the preceding cases. In fact the transfer matrices can be written down by analogy with the three preceding cases.

16. Acknowledgment

The author wishes to thank Mr. D. Petersen for his most helpful comments and for his painstaking work in reading and correcting the manuscript of this paper.

17. References

1. E. A. Guillemin, "Communication Networks", Vol. 2, Ch. IV (John Wiley, New York, 1935).
2. L. A. Pipes, "The matrix theory of four-terminal networks", *Phil. Mag.*, **30**, p. 370, 1940.
3. L. C. Peterson, "Equivalent circuits of linear active four-terminal networks", *Bell Syst. Tech. J.*, **27**, p. 593, 1948.
4. S. R. Deards, "Matrix theory applied to thermionic valve circuits", *Electronic Engineering*, **24**, p. 264, 1952.
5. M. N. Srikantaswamy and K. K. Nair, "Matrix analysis of vacuum-tube circuits", *J. Instn Telecommun. Engrs, India*, **5**, p. 23, 1958.
6. M. N. Srikantaswamy and K. K. Nair, "Analysis of active networks by admittance matrices", *J. Instn Telecommun. Engrs, India*, **5**, p. 186, 1959.
7. J. S. Bell and K. Brewster, "An introduction to matrices and their use in transistor circuits", *Electronic Engineering*, **31**, p. 98, 1959.

18. Appendix 1

The Transfer Matrix of the Four-Pole of Fig. 14

Referring to Fig. 14,

$$i_1 = i_{1A} + i_{1B} \qquad \ldots\ldots(66)$$
$$i_2 = i_{2A} + i_{2B} \qquad \ldots\ldots(67)$$

Equations (44) are

$$V_1 = a_{11} V_2 + a_{12} i_{2A} \qquad \ldots\ldots(68)$$
$$i_{1A} = a_{21} V_2 + a_{22} i_{2A} \qquad \ldots\ldots(69)$$

Equations (45) are

$$V_1 = b_{11} V_2 + b_{12} i_{2B} \qquad \ldots\ldots(70)$$
$$i_{1B} = b_{21} V_2 + b_{22} i_{2B} \qquad \ldots\ldots(71)$$

$$C = \frac{1}{\beta_{21} + \alpha_{21}\beta} \cdot \begin{pmatrix} \alpha_{11}\beta_{21} + \alpha_{21}\beta_{22} & , & \alpha_{11}\beta_{11} + \alpha_{22}\beta_{22} + \alpha_{12}\beta_{21} + \alpha_{21}\beta_{12} - \alpha\beta - 1 \\ \alpha_{21}\beta_{21} & , & \alpha_{21}\beta_{11} + \alpha_{22}\beta_{21} \end{pmatrix} \qquad \ldots\ldots(65)$$

If the feedback network is bilateral it is readily shown that eqn. (65) reduces to eqn. (61).

Eliminating i_{2A} and i_{2B} from eqns. (67), (68) and (70)

$$V_1 = \frac{a_{12}b_{11}+a_{11}b_{12}}{a_{12}+b_{12}}\cdot V_2 + \frac{a_{12}b_{12}}{a_{12}+b_{12}}\, i_2 \quad \text{...(72)}$$

provided $\quad a_{12}+b_{12} \neq 0.$

Eliminating i_{1A} and i_{1B} from eqns. (66), (69) and (71),

$$i_1 = (a_{21}+b_{21})V_2 + a_{22}i_{2A} + b_{22}i_{2B} \quad \text{......(73)}$$

Eliminating i_{2A} and i_{2B} from eqns. (68), (70) and (73),

$$a_{12}b_{12}i_1 = a_{12}b_{12}(a_{21}+b_{21})V_2 + a_{22}b_{12}(V_1-a_{11}V_2)+b_{22}a_{12}(V_1-b_{11}V_2) \quad \text{......(74)}$$

Eliminating V_1 from eqns, (72) and (74),

$$i_1 = \frac{(a_{12}+b_{12})(a_{21}+b_{21})-(a_{11}-b_{11})(a_{22}-b_{22})}{a_{12}+b_{12}}V_2 + \frac{a_{22}b_{12}+b_{22}a_{12}}{a_{12}+b_{12}}\, i_2 \quad \text{......(75)}$$

provided $\quad a_{12}+b_{12} \neq 0.$

Thus the transfer matrix of Fig. 14, is given by

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = \frac{1}{a_{12}+b_{12}} \begin{pmatrix} a_{12}b_{11}+a_{11}b_{12} & , & a_{12}b_{12} \\ (a_{12}+b_{12})(a_{21}+b_{21})-(a_{11}-b_{11})(a_{22}-b_{22}) & , & a_{22}b_{12}+a_{12}b_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \quad \text{......(76)}$$

provided $\quad a_{12}+b_{12} \neq 0.$

The determinant of the transfer matrix is readily shown to be

$$\frac{ab_{12}^2+ba_{12}^2+(a+b)a_{12}b_{12}}{(a_{12}+b_{12})^2}$$

If the two four-poles A and B have unity determinants, i.e. $a = b = 1$, then the above determinant is unity also.

19. Appendix 2

The Transfer Matrix of the Four-Pole of Fig. 17

Referring to Fig. 17,

$$i_i = i_{i\alpha} + i_{i\beta} \quad \text{......(77)}$$

$$i_o = i_{o\alpha} + i_{o\beta} \quad \text{......(78)}$$

Also,

$$\begin{pmatrix} V_i \\ i_{i\alpha} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & , & \alpha_{12} \\ \alpha_{21} & , & \alpha_{22} \end{pmatrix} \begin{pmatrix} V_o \\ i_{o\alpha} \end{pmatrix}$$

i.e.

$$V_i = \alpha_{11}V_o + \alpha_{12}i_{o\alpha} \quad \text{......(79)}$$

$$i_{i\alpha} = \alpha_{21}V_o + \alpha_{22}i_{o\alpha} \quad \text{......(80)}$$

Equations (47) are

$$V_o = \beta_{11}V_i - \beta_{12}i_{i\beta} \quad \text{......(81)}$$

$$-i_{o\beta} = \beta_{21}V_i - \beta_{22}i_{i\beta} \quad \text{......(82)}$$

Eliminating $i_{o\alpha}$ and $i_{o\beta}$ from eqns. (78), (79) and (82),

$$V_i = \alpha_{11}V_o + \alpha_{12}i_o + \alpha_{12}\beta_{21}V_i - \alpha_{12}\beta_{22}i_{i\beta} \quad \text{...(83)}$$

Eliminating $i_{i\beta}$ from eqns. (81) and (83),

$$V_i = \frac{\alpha_{11}\beta_{12}+\alpha_{12}\beta_{22}}{\beta_{12}+\alpha_{12}\cdot\beta}\cdot V_o + \frac{\alpha_{12}\beta_{12}}{\beta_{12}+\alpha_{12}\cdot\beta}\, i_o \quad \text{...(84)}$$

provided $\quad \beta_{12}+\alpha_{12}\cdot\beta \neq 0.$

Eliminating $i_{i\alpha}$ and $i_{i\beta}$ from eqns. (77), (80) and (81),

$$\beta_{12}i_i = \beta_{12}\alpha_{21}V_o + \beta_{12}\alpha_{22}i_{o\alpha} - V_o + \beta_{11}V_i \quad \text{...(85)}$$

Eliminating $i_{o\alpha}$ from eqns. (79) and (85),

$$\alpha_{12}\beta_{12}i_i = \alpha_{12}\beta_{12}\alpha_{21}V_o + \beta_{12}\alpha_{22}V_i - \beta_{12}\alpha_{22}\alpha_{11}V_o \quad \text{......(86)}$$

Eliminating V_i from eqns. (84) and (86),

$$i_i = \frac{\alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-\alpha\beta-1}{\beta_{12}+\alpha_{12}\cdot\beta}\cdot V_o + \frac{\alpha_{12}\beta_{11}+\alpha_{22}\beta_{12}}{\beta_{12}+\alpha_{12}\cdot\beta}\cdot i_o \quad \text{......(87)}$$

provided $\quad \beta_{12}+\alpha_{12}\cdot\beta \neq 0.$

Thus the transfer matrix of the system of four-poles of Fig. 17, is given by

$$\begin{pmatrix} V_i \\ i_i \end{pmatrix} = \frac{1}{\beta_{12}+\alpha_{12}\cdot\beta} \begin{pmatrix} \alpha_{11}\beta_{12}+\alpha_{12}\beta_{22} & , & \alpha_{12}\beta_{12} \\ \alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-\alpha\beta-1 & , & \alpha_{12}\beta_{11}+\alpha_{22}\beta_{12} \end{pmatrix} \begin{pmatrix} V_o \\ i_o \end{pmatrix} \quad \text{......(88)}$$

provided $\quad \beta_{12}+\alpha_{12}\cdot\beta \neq 0.$

20. Appendix 3

The Transfer Matrix of the Four-Pole of Fig. 18

Referring to Fig. 18,

$$V_1 = V_{1A} - V_{1B} \quad \ldots\ldots(89)$$

and

$$i_2 = i_{2A} + i_{2B} \quad \ldots\ldots(90)$$

Equations (52) are

$$V_{1A} = a_{11} V_2 + a_{12} i_{2A} \quad \ldots\ldots(91)$$

$$i_1 = a_{21} V_2 + a_{22} i_{2A} \quad \ldots\ldots(92)$$

Equations (53) are

$$V_{1B} = b_{11} V_2 + b_{12} i_{2B} \quad \ldots\ldots(93)$$

$$-i_1 = b_{21} V_2 + b_{22} i_{2B} \quad \ldots\ldots(94)$$

Eliminating i_{2A} and i_{2B} from eqns. (90), (92) and (94),

$$i_1 = \frac{a_{21} b_{22} + b_{21} a_{22}}{b_{22} - a_{22}} V_2 + \frac{a_{22} b_{22}}{b_{22} - a_{22}} i_2 \quad \ldots(95)$$

provided $\quad b_{22} - a_{22} \neq 0.$

Eliminating V_{1A} and V_{1B} from eqns. (89), (91) and (93),

$$V_1 = (a_{11} - b_{11}) V_2 + a_{12} i_{2A} - b_{12} i_{2B} \ldots\ldots(96)$$

Eliminating i_{2A} and i_{2B} from eqns. (92), (94) and (96),

$$a_{22} b_{22} V_1 = [(a_{11} - b_{11}) a_{22} b_{22} - a_{12} a_{21} b_{22} + b_{12} b_{21} a_{22}] V_2 + (a_{12} b_{22} + b_{12} a_{22}) i_1 \ldots(97)$$

Eliminating i_1 from eqns. (95) and (97),

$$V_1 = \frac{(a_{12} + b_{12})(a_{21} + b_{21}) - (a_{11} - b_{11})(a_{22} - b_{22})}{b_{22} - a_{22}} V_2 + \frac{a_{12} b_{22} + b_{12} a_{22}}{b_{22} - a_{22}} i_2 \ldots\ldots(98)$$

The transfer matrix of the system of Fig. 18, is given by

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = \frac{1}{b_{22} - a_{22}} \begin{pmatrix} (a_{12} + b_{12})(a_{21} + b_{21}) - (a_{11} - b_{11})(a_{22} - b_{22}) & , & a_{12} b_{22} + b_{12} a_{22} \\ a_{21} b_{22} + b_{21} a_{22} & , & a_{22} b_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \ldots\ldots(99)$$

provided $\quad b_{22} - a_{22} \neq 0.$

The determinant of the transfer matrix is readily shown to be

$$\frac{ab_{22}^2 + ba_{22}^2 - (a+b) a_{22} b_{22}}{(b_{22} - a_{22})^2}$$

If the two four-poles A and B have unity determinants, i.e. if $a = b = 1$, then the above determinant is unity also.

21. Appendix 4

The Transfer Matrix of the Four-Pole of Fig. 20

Referring to Fig. 20,

$$V_i = V_{i\alpha} - V_{i\beta} \quad \ldots\ldots(100)$$

and

$$i_o = i_{o\alpha} + i_{o\beta} \quad \ldots\ldots(101)$$

Also

$$\begin{pmatrix} V_{i\alpha} \\ i_i \end{pmatrix} = \begin{pmatrix} \alpha_{11} & , & \alpha_{12} \\ \alpha_{21} & , & \alpha_{22} \end{pmatrix} \begin{pmatrix} V_o \\ i_{o\alpha} \end{pmatrix}$$

i.e.

$$V_{i\alpha} = \alpha_{11} V_o + \alpha_{12} i_{o\alpha} \quad \ldots\ldots(102)$$

and

$$i_i = \alpha_{21} V_o + \alpha_{22} i_{o\alpha} \quad \ldots\ldots(103)$$

Equations (55) are

$$V_o = \beta_{11} V_{i\beta} + \beta_{12} i_i \quad \ldots\ldots(104)$$

and

$$-i_{o\beta} = \beta_{21} V_{i\beta} + \beta_{22} i_i \quad \ldots\ldots(105)$$

Eliminating $i_{o\alpha}$ and $i_{o\beta}$ from eqns. (101), (103) and (105),

$$i_i = \alpha_{21} V_o + \alpha_{22} i_o + \alpha_{22} \beta_{21} V_{i\beta} + \alpha_{22} \beta_{22} i_i \ldots(106)$$

Eliminating $V_{i\beta}$ from eqns. (104) and (106),

$$i_i = \frac{\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}}{\beta_{11} - \alpha_{22} \beta} V_o + \frac{\alpha_{22} \beta_{11}}{\beta_{11} - \alpha_{22} \beta} \ldots(107)$$

provided $\quad \beta_{11} - \alpha_{22} \beta \neq 0.$

Eliminating $V_{i\alpha}$ and $V_{i\beta}$ from eqns. (100), (102) and (104),

$$\beta_{11} V_i = \alpha_{11} \beta_{11} V_o + \alpha_{12} \beta_{11} i_{o\alpha} - V_o + \beta_{12} i_i \ldots(108)$$

Eliminating $i_{o\alpha}$ from eqns. (103) and (108),

$$\alpha_{22} \beta_{11} V_i = (\beta_{11} \alpha - \alpha_{22}) V_o + (\beta_{11} \alpha_{12} + \beta_{12} \alpha_{22}) i_i \quad \ldots\ldots(109)$$

Eliminating i_i from eqns. (107) and (109),

$$V_i = \frac{\alpha_{11} \beta_{11} + \alpha_{22} \beta_{22} + \alpha_{12} \beta_{21} + \alpha_{21} \beta_{12} - \alpha\beta - 1}{\beta_{11} - \alpha_{22} \beta} V_o + \frac{\alpha_{12} \beta_{11} + \alpha_{22} \beta_{12}}{\beta_{11} - \alpha_{22} \beta} . i_o \ldots\ldots(110)$$

Thus the transfer matrix of the system of four-poles of Fig. 20, is given by,

$$\begin{pmatrix} V_i \\ i_i \end{pmatrix} = \frac{1}{\beta_{11} - \alpha_{22} \beta} \begin{pmatrix} \alpha_{11} \beta_{11} + \alpha_{22} \beta_{22} + \alpha_{21} \beta_{12} + \alpha_{12} \beta_{21} - \alpha\beta - 1 & , & \alpha_{12} \beta_{11} + \alpha_{22} \beta_{12} \\ \alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} & , & \alpha_{22} \beta_{11} \end{pmatrix} \begin{pmatrix} V_o \\ i_o \end{pmatrix} \ldots\ldots(111)$$

provided $\quad \beta_{11} - \alpha_{22} \beta \neq 0.$

22. Appendix 5

The Transfer Matrix of the Four-Pole of Fig. 22

Referring to Fig. 22,

$$V_1 = V_{1A} - V_{1B} \qquad \text{......(112)}$$

and

$$V_2 = V_{2A} - V_{2B} \qquad \text{......(113)}$$

Equations (59) are,

$$V_{1A} = a_{11} V_{2A} + a_{12} i_2 \qquad \text{......(114)}$$

and

$$i_1 = a_{21} V_{2A} + a_{22} i_2 \qquad \text{......(115)}$$

Equations (60) are,

$$V_{1B} = b_{11} V_{2B} - b_{12} i_2 \qquad \text{......(116)}$$

and

$$-i_1 = b_{21} V_{2B} - b_{22} i_2 \qquad \text{......(117)}$$

Eliminating V_{2A} and V_{2B} from eqns. (113), (115) and (117),

$$i_1 = \frac{a_{21} b_{21}}{a_{21}+b_{21}} \,.\, V_2 + \frac{a_{22} b_{21} + a_{21} b_{22}}{a_{21}+b_{21}} \, i_2 \qquad \text{...(118)}$$

provided $\quad a_{21}+b_{21} \neq 0.$

Eliminating V_{1A} and V_{1B} from eqns. (112), (114) and (116),

$$V_1 = a_{11} V_{2A} - b_{11} V_{2B} + (a_{12}+b_{12}) i_2 \qquad \text{...(119)}$$

Eliminating V_{2A} and V_{2B} from eqns. (115), (117) and (119),

$$a_{21} b_{21} V_1 = (a_{11} b_{21} + a_{21} b_{11}) i_1 + [a_{21} b_{21}(a_{12}+b_{12}) - a_{11} a_{22} b_{21} - a_{21} b_{12} b_{22}] i_2 \qquad \text{......(120)}$$

Eliminating i_1 from eqns. (118) and (120),

$$V_1 = \frac{a_{11} b_{21} + a_{21} b_{11}}{a_{21}+b_{21}} \,.\, V_2 + \frac{a_{11} b_{22} + a_{22} b_{11} + a_{12} b_{21} + a_{21} b_{12} - a - b}{a_{21}+b_{21}} \,.\, i_2 \qquad \text{......(121)}$$

Thus the transfer matrix of the system of four-poles of Fig. 22, is given by,

$$\begin{pmatrix} V_1 \\ i_1 \end{pmatrix} = \frac{1}{a_{21}+b_{21}} \begin{pmatrix} a_{11} b_{21} + a_{21} b_{11} \,, & (a_{12}+b_{12})(a_{21}+b_{21}) - (a_{11}-b_{11})(a_{22}-b_{22}) \\ a_{21} b_{21} \,, & a_{22} b_{21} + a_{21} b_{22} \end{pmatrix} \begin{pmatrix} V_2 \\ i_2 \end{pmatrix} \qquad \text{...(122)}$$

provided $\quad a_{21}+b_{21} \neq 0.$

It is readily shown that the determinant of the transfer matrix is

$$\frac{ab_{21}^2 + ba_{21}^2 + (a+b) a_{21} b_{21}}{(a_{21}+b_{21})^2}$$

If the two four-poles A and B have unity determinants, i.e. if $a = b = 1$, then the above determinant is unity also.

23. Appendix 6

The Transfer Matrix of the Four-Pole of Fig. 24

Referring to Fig. 24,

$$V_i = V_{i\alpha} - V_{i\beta} \qquad \text{......(123}$$

and

$$V_o = V_{o\alpha} - V_{o\beta} \qquad \text{......(124}$$

Also from the figure,

$$i_\alpha = \alpha_{11} V_{o\alpha} + \alpha_{12} i_o \qquad \text{......(125)}$$

and

$$i_i = \alpha_{21} V_{o\alpha} + \alpha_{22} i_o \qquad \text{......(126)}$$

Equations (62) are,

$$V_{o\beta} = \beta_{11} V_{i\beta} + \beta_{12} i_i \qquad \text{......(127)}$$

and

$$i_o = \beta_{21} V_{i\beta} + \beta_{22} i_i \qquad \text{......(128)}$$

Eliminating $V_{o\alpha}$ and $V_{o\beta}$ from eqns. (124), (126) and (127),

$$i_i = \alpha_{21} V_o + \alpha_{21} \beta_{11} V_{i\beta} + \alpha_{21} \beta_{12} i_i + \alpha_{22} i_o \qquad \text{...(129)}$$

Eliminating $V_{i\beta}$ from eqns. (128) and (129),

$$i_i = \frac{\alpha_{21} \beta_{21}}{\beta_{21} + \alpha_{21} \beta} \,.\, V_o + \frac{\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}}{\beta_{21} + \alpha_{21} \beta} \,.\, i_o \qquad \text{......(130)}$$

provided $\quad \beta_{21} + \alpha_{21} \beta \neq 0.$

Eliminating $V_{i\alpha}$ and $V_{i\beta}$ from eqns. (123), (125) and (128),

$$\beta_{21} V_i = \beta_{21} \alpha_{11} V_{o\alpha} + (\beta_{21} \alpha_{12} - 1) i_o + \beta_{22} i_i \qquad \text{...(131)}$$

Eliminating $V_{o\alpha}$ from eqns. (126) and (131),

$$V_i = \frac{\alpha_{11} \beta_{12} + \alpha_{21} \beta_{22}}{\beta_{21} + \alpha_{21} \beta} \, V_o + \frac{\alpha_{11} \beta_{11} + \alpha_{22} \beta_{22} + \alpha_{12} \beta_{21} + \alpha_{21} \beta_{12} - \alpha\beta - 1}{\beta_{21} + \alpha_{21} \beta} \, i_o \qquad \text{......(132)}$$

provided $\quad \beta_{21} + \alpha_{21} \beta \neq 0.$

Thus the transfer matrix of the system of four-poles of Fig. 24, is given by

$$\begin{pmatrix} V_i \\ i_i \end{pmatrix} = \frac{1}{\beta_{21}+\alpha_{21}\beta} \begin{pmatrix} \alpha_{11}\beta_{21}+\alpha_{21}\beta_{22} & , & \alpha_{11}\beta_{11}+\alpha_{22}\beta_{22}+\alpha_{12}\beta_{21}+\alpha_{21}\beta_{12}-\alpha\beta-1 \\ \alpha_{21}\beta_{21} & , & \alpha_{21}\beta_{11}+\alpha_{22}\beta_{21} \end{pmatrix} \begin{pmatrix} V_o \\ i_o \end{pmatrix} \quad \ldots\ldots(133)$$

provided $\beta_{21}+\alpha_{21}\beta \neq 0.$

24. Appendix 7

The Determinant of the Transfer Matrix of a Passive Four-Pole is Unity

A general verification of this result is difficult, but the following work verifies the result for most passive networks encountered in practice.

The transfer matrices of a series impedance Z and a shunt admittance Y are

$$\begin{pmatrix} 1 & , & Z \\ 0 & , & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & , & 0 \\ Y & , & 1 \end{pmatrix}$$

respectively. Each has unity determinant.

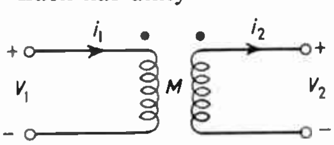


Fig. 25. A mutual inductance.

The transfer matrix of the mutual inductance shown in Fig. 25 is given by

$$V_1 = 0.V_2 - sM.i_2$$

and
$$i_1 = \frac{1}{sM}.V_2 + 0.i_2$$

where s is the complex frequency. That is, the transfer matrix of the four-pole is

$$\begin{pmatrix} 0 & , & -sM \\ \dfrac{1}{sM} & , & 0 \end{pmatrix}$$

The determinant of the transfer matrix of this passive network is unity.

If the transfer matrices of a pair of four-poles A and B have transfer matrices

$$\begin{pmatrix} a_{11} & , & a_{12} \\ a_{21} & , & a_{22} \end{pmatrix} \text{ and } \begin{pmatrix} b_{11} & , & b_{12} \\ b_{21} & , & b_{22} \end{pmatrix}$$

with determinants a and b respectively, then the transfer matrix of A and B in series cascade has determinant ab, for the rule for multiplying square matrices is the same as for multiplying determinants. Then if A and B have transfer matrices with unity determinant then the transfer matrix of the series cascaded four-poles has unity determinant also.

It has been shown in the earlier Appendices that the combination of pairs of bilateral four-poles in any arrangement leads to a system with transfer matrix of unity determinant if the constituent four-poles have transfer matrices with unity determinant.

Since most passive networks can be built up by repetitively combining four-poles in one of the above fashions starting with basic four-poles of series impedances, shunt admittances, and mutual inductances, having transfer matrices of unity determinant, it follows that such four-poles of passive elements have transfer matrices with unity determinant.

Manuscript first received by the Institution on 28th November 1960 and in final form on 30th January 1961. (Paper No. 644.)

APPLICANTS FOR ELECTION AND TRANSFER

As a result of its meeting on 25th May the Membership Committee recommended to the Council the following elections and transfers.

In accordance with a resolution of Council, and in the absence of any objections, the election and 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.

Transfer from Associate Member to Member

OAKES, Francis. *Enfield, Middlesex.*

WHITWELL, Arthur Leslie. *Glasgow.*
WRAY, Arthur George, M.A. *St. Albans, Herts.*

Transfer from Graduate to Member

EAST, Alexander Maurice. *Aveley, Essex.*

Direct Election to Associate Member

COLE, Vivian Elfed. *Waltham Cross, Hertfordshire.*

FENNER, Nigel Guy. *London, S.W.20.*

HALL, Raymond James. *Basingstoke, Hampshire.*
HAWKINS, Major William Frederick Duncan, B.Sc.(Eng.). R. Sigs. *Singapore.*
HYMANS, Anthony Jack, B.Sc., M.Sc. *Faringdon, Berkshire.*

VISONTAI, Peter. *London, N.W.6.*

Transfer from Graduate to Associate Member

DUNLOP, Edward Goodwin. *Southall, Middlesex.*

GALLAGHER-DAGGITT, George Edward. *Chelmsford, Essex.*
GARDINER, Alan. *Wells, Somerset.*

HALLIDAY, Douglas Frank. *Dumfries.*
HICKS, Charles Ernest. *Ilford, Essex.*
HOATH, William Daniel. *Wooburn Green, Bucks.*

LAKDAWALA, Homi Feroze. *Wednesfield, Staffs.*

Transfer from Student to Associate Member

SABBAH, Prosper Benjamin. *Kfar-Ata, Israel.*

Direct Election to Associate

ALEXANDER, Donald Geoffrey Morley. *Bedhampton, Hampshire.*
ALLCOCK, Michael. *Stevenage, Hertfordshire.*

BROOKER, Henry Stewart. *Norwich, Norfolk.*

Direct Election to Graduate

BARNETT, Ian James. *Great Malvern, Worcestershire.*
BARRETT, Michael. *London, S.E.14.*
BELCHER, William Leslie. *Ashford, Middlesex.*
BOOTHAM, Martin. *Pinner, Middlesex.*
BOYCE, John Patrick. *London, N.4.*

CASS, Derek Charles. *Sunbury-on-Thames, Middlesex.*

DOWNING, Raymond Meredith Maurice. *London, N.14.*

ESTALL, Geoffrey William. *Chigwell, Essex.*

FARRELLY, John Peter. *London, S.W.18.*

HOGBEN, Roger Roy. *Twickenham, Middlesex.*

MA, Ying Ping Richard, B.Sc. *Hong Kong.*

ONG, Kim Yeow. *Jesselton, North Borneo.*

POWRIE, Thomas Drew. *Stevenage, Hertfordshire.*

SAVILL, Derek Leslie. *Welwyn Garden City, Hertfordshire.*
SMITH, John. *Lytham St. Annes, Lancashire.*
STAYTON-DAVIS, Anthony Maurice. *London, W.11.*

THOMPSON, Peter Albert. *Newcastle-upon-Tyne, Northumberland.*

Transfer from Student to Graduate

ALLAN, William. *Lochwinnoch, Renfrewshire.*

HOSANGDI, Rabindranath Radhakrishna, B.Sc. *Bombay, India.*

THEWLIS, Derek. *Kingston-upon-Thames, Surrey.*

ZEPLER, Matthew Martin, B.A.(Cantab.). *Southampton, Hants.*

STUDENTSHIP REGISTRATIONS

The following students were registered at the 25th May meeting of the Committee.

AGGARWAL, Raj Kumar. *Ferozepore, India.*
ANTONY, C. V. *Singapore.*

BECKLEY, Philip, B.Sc. *Newport, Monmouthshire.*
BHATNAGAR, Vinod Kumar, B.Sc. *New Delhi, India.*
BOSE, Tapas Kumar, B.Sc. *Singapore.*
BOYD, Samuel. *Belfast, Northern Ireland.*
BRIDGER, Brian Anthony. *Enfield, Middlesex.*

DRIVER, Anthony John. *B.F.P.O.69.*

EKANEM, Ukpabid Akpan. *Lagos, Nigeria.*

FERNANDO, Warnaculasuriya Gregory Alfred. *London, S.E.4.*
FORSEY, Michael. *London, S.W.17.*

GREENWOOD, Anthony. *St. Lucia, B.W.I.*
GULATI, Amrit Singh, B.Sc. *Simla, India.*

HARDING, James John. *Brampton, Cumberland.*
HEALD, David Routledge. *Nottingham.*

ITAMBO, Ani Ukit. *Lagos, Nigeria.*

JONES, William. *Middlesbrough, Yorkshire.*

LAU, Soon Choung. *Jesselton, North Borneo.*
LEWIS, Raymond James, B.Sc. *Ammanford, South Wales.*

McKOY, Eric Halstead. *Ibadan, Nigeria.*
MABIN, Brian. *Plymouth, Devon.*
MALIK, Dharm Vir. *Jullunder Cantt, India.*
MARSHALL, Derek Lawrence, B.Sc. *London, E.13.*
MILES, Arthur Crawford. *Weston-super-Mare, Somerset.*
MOORE, Christopher Barry. *Ilford, Essex.*

OBIANWU, Emmanuel Azubuike. *Ikeja, Nigeria.*
OGU, Ignatius Dike. *Ibadan, Nigeria.*
ONWUKWE, Benson Chima. *Lagos, Nigeria.*

PARKINSON, Harry. *Bolton, Lancashire.*
PAVILONIS, John. *Airdrie, Lanarkshire.*

RAHMANI, Ghulam, B.Sc. *London, S.E.5.*
RIZVI, Syed Mohammed Mustahsan. *Mississippi, U.S.A.*
ROBERTSON, Lewis Myles. *Kirkwall, Orkney Islands.*

SINNECKER, Alexander George. *Kowloon, Hong Kong.*
SMITH, Eric. *Reading, Berkshire.*
STEPHENS, Richard Frederick. *Knutsford, Cheshire.*
SYMONS, John Gilbert. *Cardiff, Glamorgan.*

TEEKARAM, Krishnasami. *Bangalore, India.*
THOMSEN, Roy Dudley. *Marandellas S. Rhodesia.*
TONGUE, Brian Leonard. *Cheshunt, Hertfordshire.*

VYAS, Dinker Shambhushankar. *London, S.W.16.*

WATSON, Brian David. *Harlow, Essex.*
WIFFILL, Edward John James. *Plymouth, Devon.*

U.D.C. 535.64: 621.397.132

"Land" Colour—A Discussion

Report of the Proceedings of a Discussion on Two-Colour Projection with particular reference to Applications to Colour Television

Land's System of Two-Colour Projection

By

M. H. WILSON†

AND

R. W. BROCKLEBANK†

Presented at an Institution meeting held in London on 28th September 1960.

Summary: A brief introduction is given of the physical principles involved in two-colour reproduction. The limitations in reproduction of various colours are discussed.

Land's system of two-colour projection[1] is essentially a reduction from the method used for the very first colour photograph, demonstrated by Maxwell a hundred years ago. Maxwell took three photographs of a scene through red, green and blue filters each passing about one-third of the spectrum, and then projected the three positive transparencies through the same three filters so that the three images were in register on the screen, the colours combining additively to reproduce the original scene. Such a system of analysis in terms of three primary colours followed by additive recombination is colorimetrically equivalent to the processes used in colour television.

Land's variant makes two reductions: the first is to dispense with the blue separation altogether, thus ignoring the contribution of short-wave radiation to the visual sense. A two-colour picture thus projected in red and green light gives a result equivalent to that obtained by looking at the scene through a yellow glass or illuminated by yellow light, in which short-wave radiation is absent or cut out. Land's second reduction is to remove the green filter from the projecting lantern; the photograph is still exposed through a green filter, so it is a normal green separation positive, but projected in unfiltered light, generally called "white light", though actually the distinctly yellowish light characteristic of incandescent lamps. The red separation is still projected in red light, so that the combined total illumination is a reddish yellow or yellowish pink.

In spite of the reductions, pictures projected by Land's method may still show a range of colours including not only red, white and pink, but orange, brown, purple, grey, green, blue and yellow, in the sense that these are perceived as the colours of objects depicted in the scene. Various authors[2-5] have shown how these results can be explained in terms of known principles of colour vision, such as induction, adaptation, small-field tritanopia, and so on, and indeed the main importance of Land's work seems to lie in the field of demonstrating certain characteristics of colour perception. But the question inevitably arises as to whether this method could provide a simpler but still adequate substitute for three-colour methods.

Land's pictures all show interior scenes, either portraits or still-life table-top settings of carefully selected objects. To test the adequacy of colour reproduction by Land's method, a more suitable scene would include samples showing each principal hue at a number of different lightnesses and saturations, exhibiting clearly the three independent variables used in describing colour. Pages from the Munsell Book of Colour are appropriate for this purpose. When reproduced by Land's method there is no difference in hue between the orange-red page and the purple-red page, though on each the lighter samples might be more orange, even slightly yellow, and the darker samples more purple. Similarly the green page and the blue page both show an identical range of colours in the Land-type projection, both showing more blue in the darker samples and more green in the brighter. Thus hue is no longer an independent variable, but depends partly on lightness; the yellow and blue that are perceived do not depend on the yellowness or blueness of the original sample. In a two-colour projection, with two given primaries, the only variable factors controlling the colours in the final picture are the densities in the photographic positives. There are thus only two independent variables to control the colour, and no amount of photographic (or electronic) subtlety can make a two-dimensional range adequately represent a three-

† Goethean Science Foundation, Clent, Stourbridge, Worcestershire.

dimensional range, nor provide from the evidence of the red and green separations the information given by the missing blue separation. The only way of making a two-colour representation of a scene appear at all realistic is to limit the range of colours in the original scene to a two-dimensional array; apart from the device of selecting the objects portrayed, this can be done most effectively by using a yellow illuminant, and in order for this to appear natural it has to represent electric or candle light as used for much interior lighting.

The reason for the appearance of neutral colours, and greens and blue-greens out of a mixture of red and yellowish lights is due to adaptation to the apparent illuminant, and contrast induced by the strong reds. While this adds to the magic of the demonstration it does not improve the accuracy or faithfulness of the reproduction. In fact a rather better range of colours is available if the primary projecting lights are both coloured. Fully saturated red and green primaries, as used by Maxwell, give a very intensely coloured picture in which the overall yellowness is too marked; white objects are unmistakably yellow. Using complementary colours for the primaries, for instance red and cyan, the combined light, and hence the apparent illuminant, is actually white, and white objects are truly white; but the obvious absence of blueness and yellowness is sharply contrasted with the strong reds and cyans, giving a very artificial appearance to the picture. The best compromise seems to be to use an orange-red and a slightly bluish green, neither very strongly saturated; the combined light is rather yellowish, not too strong to spoil the sense of whiteness in white objects, but sufficient to bring a greater diversity of hue into the scene—yellowness in light objects and complementary blueness in dark ones. The possibility of perceiving yellow and blue is further enhanced if these colours are confined to small patches, and if the total illumination is kept low (that is, distinctly dim) besides being biased towards the yellow side.

If the scene to be portrayed contains the normal range of variation found in coloured objects, then even the best two-primary system is bound to produce failures and ambiguities; there are bound to be colours that look different in the original but which are indistinguishable in the reproduction. If it is the contribution made by the short-wave radiations to vision that is eliminated, as in Land's system, then the reproduction will fail to distinguish violet or deep blue from black, purple from brown, mauve from orange, pink from amber, grey from olive green, bluish green from yellow-green, white from light yellow, and so on. The colour that appears may happen to be the right one, it may be near enough to be acceptable to an uncritical observer, or it may be hopelessly wrong and make nonsense of the scene. Even where a picture is tolerable in a two-colour version, the addition of the missing third component at once resolves the ambiguity of the colours perceived and shows up the defects inherent in the reduced system.

It has been asserted[6] that the early use of two-primary processes for colour cinematography, with their inevitable distortions and inadequacies, produced a widespread prejudice against any sort of colour in films, a prejudice that has not yet completely died out. It would be a pity if the same were to happen in colour television.

References

1. E. H. Land, "Color vision and the natural image," Part I, *Proc. Natl. Acad. Sci. U.S.*, **45**, pp. 115–29, January 1959; Part II, *ibid*, **45**, pp. 636–44, April 1959.

   "Experiments in color vision", *Scientific American*, **200**, pp. 84–99, May 1959, *ibid*, **201**, p. 16, September 1959.

2. A. Karp, "Colour-image synthesis with two unorthodox primaries", *Nature*, **184**, pp. 710–12, 29th August 1959.

   "Tritanopia and two-colour image synthesis", *Nature*, **188**, pp. 40–2, 1st October 1960.

3. G. L. Walls, "Land! Land!", *Psychological Bulletin*, **57**, pp. 29–48, January 1960.

4. D. B. Judd, "Appraisal of Land's work on two-primary color projections", *J. Opt. Soc. Amer.*, **50**, pp. 254–68, March 1960; reprinted in *J. Phot. Sci.*, **8**, pp. 125–40, July–August 1960.

5. M. H. Wilson and R. W. Brocklebank, "Two-colour projection phenomena", *J. Phot. Sci.*, **8**, pp. 141–50, July–August 1960.

6. A. Cornwell-Clyne, "Colour Cinematography", pp. 661–2, 3rd edn (Chapman & Hall, London, 1951).

Manuscript received by the Institution on 19th December 1960. (*Contribution No.* 9.)

This opening contribution to the discussion meeting was supported by demonstrations of two and three colour reproductions.

U.D.C. 535.64: 621.397.132

The Range of Colours Excited by a Two-Colour Reproduction System

By

W. N. SPROSON, M.A.†

Presented at an Institution meeting held in London on 28th September 1960

Summary: Some experiments are described in which two-colour and three-colour reproductions of the same scene are compared. Optical projection of the separation positives was used in the main but a description of a colour television demonstration is also given. The accuracy of a two-colour process is assessed on the basis of the colour names given to specified areas of the colour reproductions in the two-colour and three-colour versions.

1. Introduction

In the early days of colour cinematography a number of two-colour processes[1] were operated with a fair measure of success. One such process was the Raycol additive two-colour process, which operated during 1928–29. Red and green separation positives were projected and combined in register on the screen. Originally the green separation was projected through a blue-green filter but it was found that the filter was unnecessary. The red positive was projected through a red filter. The effect is described by Clyne[1] as being "astonishingly good". Baird experimented with a two-colour television system and gave a demonstration of a 600-line picture in 1941 using a two-colour frame sequential system.[2] More recently interest in this subject has been stimulated by Land,[3] who has claimed that the "full gamut of color" can be produced by a two-colour process.

One form of two-colour process described by Land closely resembles the Raycol process described above. The green and red separations are used; the green is projected in a black-and-white projector and the red separation is projected through a tricolour red filter. Land has not claimed that the colours are accurately reproduced, but he states that pleasantly coloured scenes are generated by these means. The experiments about to be described probably support the latter claim, but the primary purpose was to obtain statistical data on the subjective accuracy of colour reproduction of this two-colour process.

2. Experimental Procedure

Three-colour separation positives of a number of scenes were prepared; some were obtained by direct photography through tricolour red, green and blue filters. Others were obtained from Ektachrome transparencies using the recommended techniques. A triple projector (Fig. 1) was improvised by using three Aldis projectors. This arrangement, using a titanium dioxide mirror, permitted exact registration to be obtained as between the red and green images. The blue image suffered slight keystone distortion compared with the other images, but this did not appear to cause any appreciable trouble, presumably because of the relatively low luminosity of the blue component. The colour picture was projected on to a matt white screen. The size of image was about 15 in × 20 in (38 × 50 cm) and the image was viewed at a distance of about 7 ft (2·1 m). Thus the ratio of viewing distance to picture height was 5·6, which is typical of critical viewing of television pictures although it is a smaller ratio than that generally applying to the viewing of a 21 in (53 cm) television picture (405-line) in the home.

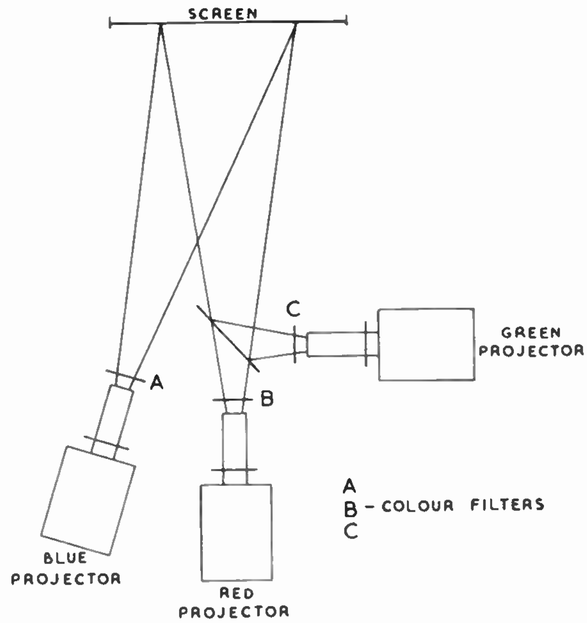


Fig. 1. Schematic of triple projector.

† The British Broadcasting Corporation, Research Department, Kingswood Warren, Tadworth, Surrey.

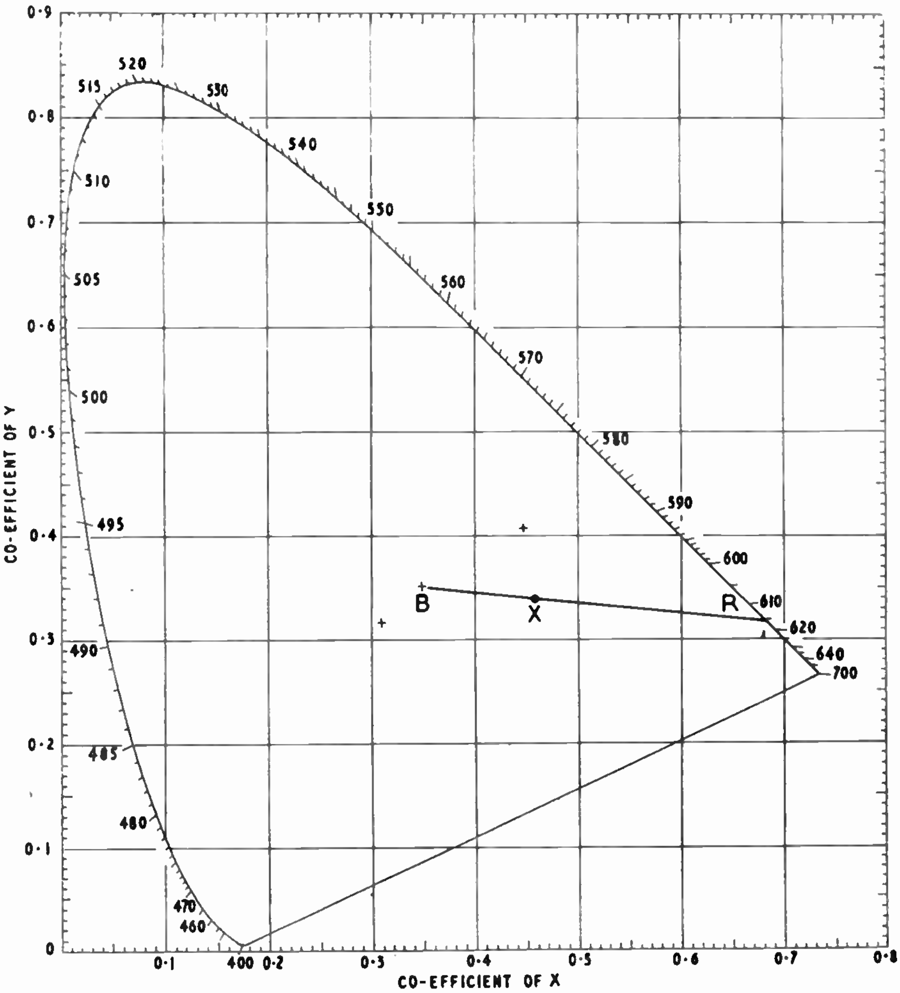


Fig. 2. Range of chromaticities objectively produced by two-colour process.

The peak-white brightness was of the order of 3 to 5 ft-lamberts (30 to 50 asb) which is rather less than that available from a colour television tube (e.g. R.C.A. type 21AXP22).

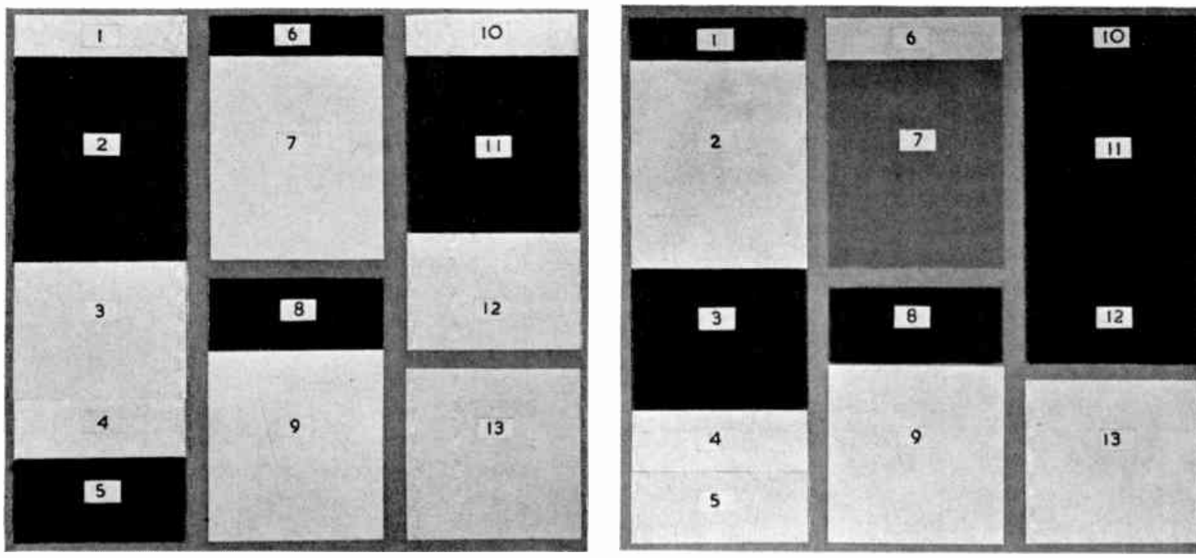
For a three-colour presentation, all three projectors were used: the projectors giving the blue and red images were fed from continuously variable voltage transformers. This feature enabled the colour balance to be varied over wide limits. The values of voltage were chosen to give a satisfactory appearance of the colour scene.

For a two-colour presentation, the blue projector was switched off and the green filter was removed from the "green" projector and replaced by a colour-temperature-raising filter.† The filter in the red projector was unaltered and the voltage on this projector was adjusted to give the best results. This implied, in practice, the adjustment for which the maximum range of colours was produced without making the whites in the two-colour reproduction obviously pink. An increase in the light flux from the red projector beyond the condition just described gave more saturated colours but the whole colour balance clearly suffered from an excess of the red component.

From this description of the synthesis of a two-colour picture, it will be observed that the image is built up by the addition of white and red light in various proportions. Hence, one might even regard this synthesis as "white plus one colour", reserving the description of "two-colour" for combinations of complementary colours such as red and cyan. The white-plus-red system will, however, be described as "two-colour".

If one ignored effects such as simultaneous contrast, it might be thought that the two-colour presentation

† This was estimated to raise the colour temperature of the projector light from 3000° K to about 4500° K. This is nearer to the chromaticity of a monochrome television tube and was found to improve the saturation of the colours given by the two-colour process.



(*a*) Red separation positive.

(*b*) Green separation positive.

Fig. 3. Colour patches.

just described would produce a very restricted range of colours which would include only reds, pinks, whites and greys. Because of simultaneous contrast effects, however, a much greater range of colours is produced and blue-greens are clearly in evidence. The mechanism by which this takes place may be as follows: the highest luminances in the scene are usually accepted as white by the observer, provided that they are not markedly coloured as compared with other objects in the scene. This means that the subjective white point moves to a position on the colour triangle somewhere between the chromaticity of illuminant B and that of tricolour red, e.g. the point X in Fig. 2. Thus from X to R we have physical stimuli which give rise to sensations of red of various intensities, whilst from X to B the sensation of the colour complementary to red, namely blue-green, is produced.†

Further discussion of the colours produced by a two-colour system will be deferred until the experimental results have been given.

3. Results

Four scenes were used as subjects for experiment. In three of these the colour separation positives were obtained by direct photography. The other picture was originally photographed on Ektachrome sheet film and the colour separations were obtained from the colour transparency.

† Since the completion of this work, a confirmation and extension of this idea has been published by A. Karp, "Colour-image synthesis with two unorthodox primaries", *Nature*, **184**, pp. 710–712, 29th August 1959.

The observers were shown a colour picture (either two- or three-colour) and asked to name the colours in certain specific areas. No restrictions were placed on the colour namings, although it was explained that consistency of naming was desirable. Having listed the colours seen in (say) the two-colour reproduction, the observer was then asked to give colour names to the same areas of picture in the three-colour reproduction. For the first three pictures six male observers were used; for the fourth picture six female and six male observers were used. Only one observer was questioned at any one time, so that the colour namings are individual opinions and not the result of a group discussion.

3.1. *Colour Patches*

This "picture" consists of thirteen colour patches of rectangular shape. Figures 3 (*a*) and 3 (*b*) show black-and-white prints of the red and green separation positives, respectively. Table 1 gives the percentage of observers who gave

$$\left.\begin{array}{l}\text{(a) identical namings}\\ \text{(b) identical and similar‡ namings}\end{array}\right\}\text{ to the two-colour and three-colour versions.}$$

In all cases the two-colour version was shown first and then the three-colour version.

‡ The definition of the "similar" class of namings is somewhat flexible and is not capable of being precisely defined, but the general principle followed was to allow slight changes of hue or saturation as being in the "similar" classification. A marked change in either quality was sufficient to make the namings "different". The only category about which there can be no argument is the "identical" one. See Appendix for the detailed results for the first picture.

Fig. 4 (*a*) Teapot slide: red separation positive.

Fig. 4 (*b*) Teapot slide: green separation positive.

Table 1

Results of observations on colour patches (Fig. 3)

Colour patch number	Colour of original	Approximate chromaticity co-ordinates of original illuminated by source C: x	y	Identical namings %	Identical and similar namings %
1	Red	0·43	0·29	50	50
2	Cyan	0·23	0·25	17	50
3	Brown	0·47	0·28	50	83
4	Orange-yellow	0·47	0·44	0	17
5	Green	0·28	0·45	0	17
6	Green	0·28	0·45	0	17
7	Red	0·43	0·29	50	67
8	Blue	0·21	0·17	0	0
9	Orange-yellow	0·47	0·44	17	17
10	Brown	0·47	0·28	67	100
11	Violet	0·23	0·15	0	0
12	Brown	0·47	0·28	50	100
13	Light green	0·36	0·49	0	0

Table 2

Results of observations on "teapot slide" (Fig. 4)

Colour patch number	Subject	Approximate colour of original	Identical namings %	Identical and similar namings %
1	Kodak data book	Red	50	83
2	"Under Milk Wood"	Cyan	0	33
3	Kodak syphon box	Yellow	0	0
4	Slide box	Red	50	83
5 (i)	P } Letters on lid of "Polly" box	Green	33	83
(ii)	O	Yellow	83	83
(iii)	L	Red	83	100
(iv)	L	Blue	67	67
(v)	Y	Yellow	67	67
6	B.B.C. notebook	Blue	0	67
7 (i)	Duster	Light green	33	33
(ii)		Green	17	17
8 (i)	Tartan	Red	50	83
(ii)		White	80	80
(iii)		Dark green	67	83
9	Notepaper on which teapot is standing	Pink	67	100

From Table 1 it is fairly obvious that certain colours are not well reproduced, at least under the circumstances that apply in Figs. 3 (*a*) and 3 (*b*), and these are orange-yellow, blue, violet and light green. Colour patch No. 2 (cyan) is moderately well reproduced with 50% of the observers in the "identical and similar" class of namings.

3.2. *Teapot Slide*

The colour separation positives were prepared from an Ektachrome transparency. This picture, Figs. 4 (*a*) and 4 (*b*), was used considerably during the field trials of colour television[4] and is an attractive composition (in contradistinction to the previous "picture"). The observers were asked to name sixteen coloured patches; some of these colours subtended large † solid angles at the observers' eyes but others did not. The results (Table 2) show that whereas a large-area yellow (such as the Kodak box) is not well reproduced in the two-colour version, nevertheless the second and fifth letters of "Polly", which are also yellow, are tolerably well reproduced. This suggests that the range of colours which can be simulated by a two-colour process is not precisely defined, but is influenced by size and also by the presence of other colours in the immediate vicinity. Likewise small-area blue (colour patch No. 5 (iv) in Table 2) is seen

† Large compared with the limit of visual acuity (1 minute of arc).

by four of the six observers, although the large-area blue in the first picture (Table 1, colour patch No. 8) was not recognizably reproduced at all.

In a picture such as the teapot slide one is surprised by the range of colours which are reproduced by a two-colour process. Nevertheless, it is difficult to regard the reproduction as satisfactory when a large bright yellow box is reproduced as pink.

Table 3

Results of observations on the portrait slide

Colour patch number	Subject	Approximate colour of original	Identical namings %	Identical and similar namings %
1	Flesh tones on face	Pink	50	100
2	Lips	Red	17	100
3	Eye colour	Grey-green	67	83
4(*a*)	Beads	Red-brown	17	33
(*b*)		Pink	50	83
5	Earring	Red	33	67
6(*a*)	Hair	Brown	50	100
(*b*)		Auburn	20	40
7(*a*)	Dress material	Green	0	33
(*b*)		Black	100	100
(*c*)		White	50	83
(*d*)		Red	50	83

3.3. *Portrait*

The separation negatives for this picture were taken directly by successive exposures through tricolour red, green and blue filters. This required the model to be still for 55 seconds, but the registration of the positives was nevertheless good, Figs. 5 (*a*) and 5 (*b*). Twelve colours were named by the observers; the results are given in Table 3.

On the whole, the two-colour version of this slide succeeds fairly well; the most obvious deficiency is the green of the dress material (colour patch No. 7 (*a*), Table 3). One point of interest is that many of the observers thought that the two-colour version was preferable because of an imperfection in colour rendering in the three-colour version, which caused certain highlights on the forehead to be reproduced with a greenish tinge. Since this was clearly an artifact it would not be fair to infer that, in general, a two-colour version would be preferred. It should, however, be borne in mind that the balance of a three-colour print can be disturbed easily, and that a trend towards either green or magenta in flesh tones is highly undesirable.

3.4. *Detergent Packets*

Six packets of detergent were photographed and direct separations made. These packets are brightly coloured, and the colours are sufficiently well known for the observers to be asked about the accuracy of

(*a*) Red separation positive. (*b*) Green separation positive.

Fig. 5. Portrait.

colour in the two-colour and three-colour versions in addition to the colour namings. Sixteen colours were named by the observers, who for this picture included six women in addition to the six men. Three of the women were shown the three-colour version first, and then the two-colour version, and the other three women saw the two-colour version first. This procedure was also followed for the six male observers. Although it is not thought that the order of showing the versions has any appreciable effect on the colour namings, this procedure should reduce the influence of any factors which depend on the order of showing.

The results of the naming of sixteen colours are given in Table 4. The range of colour names used by the women is greater than that used by the men and it will be observed that the corresponding percentages in either the "identical" class or the "identical and similar" class is usually lower for the women. Women are, in general, more interested in colour than men.

Table 4

Results of observations on detergent packet slide

Colour patch number	Colour of original	Approximate chromaticity co-ordinates of original illuminated by source C x	y	Male observers: Identical namings %	Male observers: Identical and similar namings %	Female observers: Identical namings %	Female observers: Identical and similar namings %
1	Red	0·52	0·32	50	100	50	83
2	Blue	0·20	0·20	0	0	0	0
3	Red-orange	0·53	0·34	33	100	0	67
4	Orange-yellow	0·46	0·37	0	0	0	0
5	Red	0·53	0·34	50	100	0	67
6	Blue	0·21	0·24	33	50	0	33
7	Red	0·53	0·34	67	100	33	83
8	Yellow	0·44	0·49	0	0	0	0
9	Red	0·52	0·32	50	100	0	83
10	Yellow	0·44	0·49	0	0	0	0
11	Blue	0·14	0·18	0	0	0	0
12	Red	0·52	0·32	50	83	0	50
13	Blue	0·21	0·24	0	17	0	17
14	Red	0·53	0·34	83	100	0	83
15	Yellow	0·44	0·49	0	0	0	0
16	Green	0·28	0·44	0	0	0	0

Certain colours fail almost completely in the two-colour version (e.g. blue, yellow and green) and these colours are frequently used in display and packaging. If it were desired to transmit a colour television picture similar to this slide, a two-colour process would leave a great deal to be desired in so far as colour accuracy was concerned. In reply to a question as to which of the two presentations was the more truthful, all twelve observers replied in favour of the three-colour version, including those who said that they were not very familiar with these particular packets; they could not imagine any manufacturer using the colours which are generated by the two-colour process. The only criticism made about the three-colour version was that the colours were a little too saturated. Most observers thought that the hues were correct.

3.5. *Summary of Information from the Four Pictures*

The mean percentages averaged over all colours for the "identical" and "identical and similar" classes of namings for the four slides are given in Table 5.

Table 5

Summary of observation results

Picture	Title	Mean identical naming %	Mean identical and similar naming %
1	Colour patches	23	40
2	Teapot slide	47	66
3	Portrait	42	75
4	Detergents (men)	26	47
	(women)	5	35

The most successful slides appear to be the portrait and the teapot slide, where two-thirds to three-quarters of the colours receive names which are similar, if not identical, to the namings given to a three-colour reproduction of the same slide. In the case of slides 1 and 4, the two-colour reproduction has very glaring deficiencies.

An attempt has been made to give a rough estimate of the range of chromaticities over which the two-colour reproduction operates with an approximation to accuracy. The original colours in pictures 1 and 4 were compared with the chips in the Munsell Colour Atlas and the best available matches were found. The chromaticity co-ordinates of these colour chips illuminated in daylight (illuminant C) were read off from the data published by Kelly, Gibson, and Nickerson.[5] These figures are quoted in Tables 1 and 4. Figure 6 is a chromaticity diagram showing the percentage of observers giving "identical and similar" colour names and, on the basis of these few points, an area has been drawn on the chromaticity diagram over which 50% or more of the observers use the "identical" or "similar" namings. A further contour is shown which marks an area over which virtually no accuracy in colour reproduction is achieved in

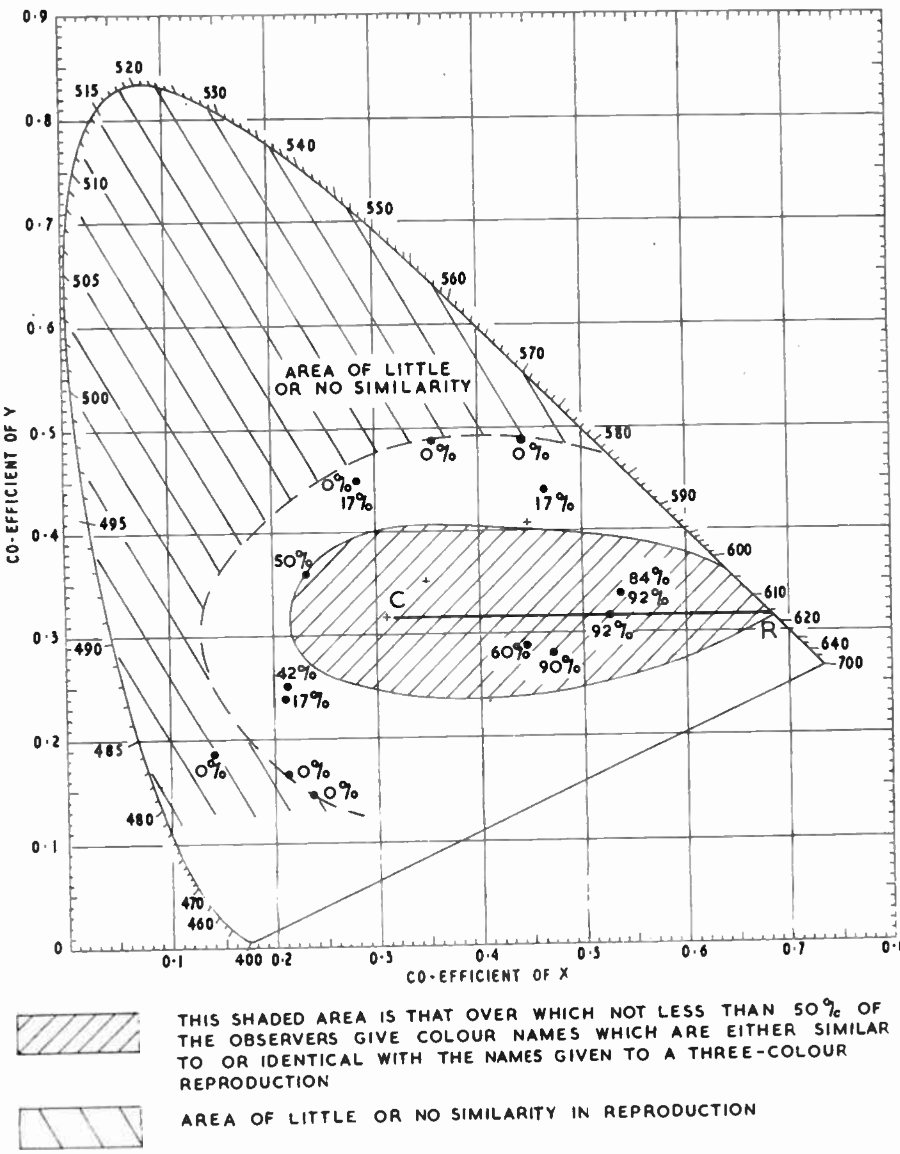


Fig. 6. Approximate accuracy of colour rendering of two-colour process.

large and medium sized areas of colour on the original article. The "objective" range of colours produced by a mixture of tricolour red and illuminant C is shown by the straight line CR (Fig. 6). The use of illuminant C in Fig. 6 is undesirable, but the available data[5] do not quote chromaticities in illuminant B. A subsidiary experiment showed little change in the range of colours produced when illuminant C is used for the green positive in place of illuminant B (although the change from illuminant A to illuminant B does seem to produce a marked improvement in the saturation). Hence, the areas shown in Fig. 6 are thought to be approximately correct, although the experimental data relate more nearly to illuminant B. In any case, far more colours would have to be investigated if any attempt were to be made to draw an accurate version of Fig. 6. Nevertheless, it is considered that this diagram represents to a first approximation the behaviour of a two-colour system according to Land in reproducing large and medium sized areas of colour.†

In smaller areas it would appear that the range of colours which can be simulated is appreciably greater than the range shown in Fig. 6. Thus yellows, greens and blues have been named in the course of these experiments, although usually these colours are either

† In interpreting Fig. 6 it should be remembered that the C.I.E. diagram is not a uniform chromaticity chart and that the green area is much exaggerated. Note that the chromaticity co-ordinates of only one of the colour patches of pictures 1 and 4 are well inside the area of little or no similarity.

desaturated or, alternatively, suffer from the admixture of neutral grey. On no occasion has a purple been reported and it seems that this colour is not simulated[6] by the particular version of two-colour reproduction which has been investigated in the present experiments.

4. Colour Television Demonstration

On 24th July 1959, a demonstration of a two-colour synthesis was given at Television Studios, Lime Grove. The picture-originating equipments were the B.B.C. Research Department film and slide scanner[7] and the Cintel 35 mm colour scanner. The R, G and B signals were fed both to a tricolour monitor[8] using an R.C.A. 21AXP22 tube and also to a trinoscope.[9] The latter display device had been modified by substituting a monochrome tube for the green cathode-ray tube and by disconnecting the blue tube. Thus, the trinoscope became identical in principle to the two-colour projection of white and red images as described in the previous sections.

A number of 3¼-in (8·3 cm) square colour transparencies were shown both in the two-colour and three-colour versions, and it was observed that the effects produced were virtually identical with those observed in the optical projection experiments. The teapot slide (Section 3.2) was amongst those used and the same faults were observed (the most glaring of which was the inability of the two-colour version to reproduce the yellow Kodak box). One particular slide, showing a green B.B.C. field strength van in rural surroundings (grass and trees), completely failed to give any colour whatsoever in the two-colour version.

A 35 mm colour film was also shown and this frequently gave rise to scenes with almost no colour in the two-colour version, although a glance at the tricolour display showed a reasonable range of colours. The only obvious merit of the two-colour synthesis was its ability to produce acceptable flesh tones. In other respects it showed considerable shortcomings amounting either to the virtual absence of colour or, on other occasions, to a complete falsification of the colours.

5. Discussion

The two-colour system (white plus red) frequently gives pleasantly coloured pictures which cover a greater range of colours than might be expected on the basis of objective colorimetry. Nevertheless, the range is so deficient (Fig. 6) that its use is not to be recommended for any system of colour television where a prime requirement is the ability to reproduce a wide range of colours (including greens, blues, yellows and purples) with a good approximation to the true colour. Standard three-colour systems at present in use in both colour photography and colour television succeed in doing this when sufficient care is taken to ensure correct colour balancing. The two-colour system is much less critical on colour balance and, in the case of flesh tones, it is very unlikely to generate unpleasant flesh tones even when the balance is disturbed. Three-colour reproduction has no such immunity and it is true that considerable care in instrumentation must be taken to obtain correct colour rendering, particularly in the vicinity of the white point. However, three-colour processes are capable of reproducing a very wide range of colours: the two-colour system investigated in the present tests has shown marked deficiencies and the experimental results do not confirm the more extreme statements made by or on behalf of Dr. Land.[3]

6. Acknowledgments

The colour television demonstration described in Section 4 was given by the operations and maintenance staff of the Colour Telecine Section of the Lime Grove studios. This work was supervised by the Designs Department of the B.B.C.

The author also wishes to thank the Director of Engineering of the B.B.C. for permission to publish this paper.

7. References

1. A. Cornwall-Clyne, "Colour Cinematography", 3rd edn., pp. 260–61, 264–68 (Chapman and Hall, London, 1951).
2. "Baird high-definition colour television", *J. Television Soc.*, 3, No. 6, pp. 171–74, 1941.
3. (*a*) E. H. Land, "Colour vision and the natural image", Part I, *Proc. National Acad. Sciences*, **45**, pp. 115–29, 1959.
   (*b*) E. H. Land, "Colour vision and the natural image", Part II, *Proc. National Acad. Sciences*, **45**, pp. 636–44, 1959.
   (*c*) E. H. Land, "Experiments in colour vision", *Scientific American*, pp. 84–99, 1959.
   (*d*) Francis Bello, "An astonishing new theory of colour", *Fortune*, pp. 144–48, 195–96, 205, May 1959.
4. W. N. Sproson, S. N. Watson and M. Campbell, "The B.B.C. Colour Television Tests: An Appraisal of Results", B.B.C. Monograph 18, May 1958.
5. K. L. Kelly, K. S. Gibson and D. Nickerson, "Tristimulus specification of the Munsell Book of Color from spectrophotometric measurements", *J. Opt. Soc. Amer.*, **33**, pp. 355–76, July 1943.
6. E. H. Land, Ref. 3 (*b*), Fig. 3.
7. Ref. 4, p. 8.
8. Ref. 4, p. 9, para. 3.5.2.
9. Ref. 4, p. 10, para. 3.5.3.

Manuscript first received by the Institution on 6th June, 1960, and in final form on 1st November, 1960. (*Paper No. 645.*)

8. Appendix

Detailed Results of Namings for Colour Patches (Section 3.1)

Colour patch number	THREE-COLOUR						TWO-COLOUR						Number of namings in the stated category		
	Observer						Observer								
	1	2	3	4	5	6	1	2	3	4	5	6	Identical	Similar	Different
1	red	light brown	desaturated magenta	red	dark pink	red	red	light brown	orange-brown	light brown	orange	red	3	0	3
2	cyan	green-blue	cyan	cyan	light blue	cyan	blue-cyan	blue-grey	desaturated cyan	blue-grey	light blue	blue	1	2	3
3	magenta-brown	medium brown	brown	brown	brown	brown	deep magenta	medium brown	chocolate	dark red	brown	brown	3	2	1
4	orange-yellow	dark yellow	desaturated yellow	yellow	light orange	orange	orange	pinkish grey	pink	pink	pink	pink	0	1	5
5	green	dark green	green	green	green	green	green-cyan	dark blue-grey	dark desaturated cyan	grey	light grey	blue	0	1	5
6	green	dark green	green	green	green	green	green-cyan	dark blue-grey	dark desaturated cyan	grey	light grey	blue	0	1	5
7	red	pink	desaturated red	red	dark pink	light red	red	brick red	desaturated red	red	scarlet	red	3	1	2
8	blue	dark blue	violet	blue	navy blue	blue	black	very dark grey	very dark desaturated green	black	black	black	0	0	6
9	orange-yellow	dark yellow	desaturated yellow	yellow-orange	light orange	orange	orange-yellow	pinkish fawn	light pink	pink	pink	pink	1	0	5
10	magenta-brown	medium brown	brown	brown	brown	brown	deep magenta	medium brown	chocolate	brown	brown	brown	4	2	0
11	blue	very dark blue	violet	blue	navy blue	dark blue	black	very dark grey	very dark desaturated green	black	black	black	0	0	6
12	magenta-brown	medium brown	brown	brown	brown	brown	deep magenta	red/brown	chocolate	brown	brown	brown	3	3	0
13	light green	light green	light yellow-green	yellow-green	light green	light green	grey	light blue-grey	light grey	light grey	light blue	light grey	0	0	6

DISCUSSION

Under the Chairmanship of Mr. V. J. Cooper (Member)

Mr. L. H. Bedford (*Past President*): I suggest that the television engineers in the audience will have been watching these extremely interesting and able demonstrations with one particular thought in mind, namely do the phenomena here exhibited offer the chance of a colour television system requiring two channels instead of three? In so far as it is possible to reach a conclusion on the basis of one demonstration, the answer to this question would appear to be in the negative; though should it be shown that a reduction to two channels would introduce an overwhelming simplification into either the complete system or more especially into the problem of the receiving tube, the question might still be regarded as up for consideration.

As a minor point, colour television engineers have now grown accustomed to certain standard series of test slides, in particular the N.T.S.C. series produced by Kodak, and it would be most helpful to them if future demonstrations could incorporate these slides.

Messrs. M. H. Wilson and R. W. Brocklebank (*in reply*): It might well be an advantage to use familiar material, and essential for any comparative tests to use identical subject matter.

Mr. J. W. L. Dixon: The results of various experiments that I have made in connection with the Land colour theory lead me to the following conclusions.

The Land system in its present form cannot produce an *accurate* full colour reproduction equal to existing orthodox three-colour systems. However, the Land system can produce very acceptable colour pictures when compared with the highly variable and more popular *photographic* colour processes.

I am not convinced that the explanation of a shift in neutrals as demonstrated by the 90 deg rotation grey scales with "white" surrounds is the complete answer to the Land phenomena. It is interesting to note that the 90 deg rotation grey scales show no sign of blue or green when viewed without the surrounding reference "white" (object/aperture modes). I have produced very acceptable Land reproduction with no reference "white" in the picture area or surround.

Messrs. Wilson and Brocklebank (*in reply*): Acceptability depends on the critical standards applied. A two-colour method *can* produce a *perfect* reproduction of a suitable scene, but will not do so for all the types of scene required.

There is no single complete answer to Land's phenomena; adaptation to the prevailing colour of the scene (the illuminant) is an important effect, and it does *not* depend on there being a particular "white" reference —the overall colour bias is what matters. Of course, *pictures* are necessarily perceived in object mode, not in aperture mode.

Mr. C. A. Bird (*Associate*): If economic conditions should produce the situation where the United Kingdom must decide between a colour system which is simple and tolerably acceptable to the average non-technical viewer, or no colour system at all, the Land System might become worthy of consideration.

The demonstration seemed to indicate that there is considerably more tolerance in the relative amplitudes of the red and white channels than could be tolerated with a tri-colour system. Would it be possible to take advantage of this fact in a television system by using a white channel with resolution comparable to the present standards, plus a red channel with restricted amplitude frequency characteristics? If such an arrangement is capable of giving any results at all, I feel it merits further study, as it would have overwhelming advantages from the point of view of bandwidth and ease of adjustment at the receiving end.

Have the authors tried simulating this condition by optical means, e.g. by using a white slide with full detail, plus a "fuzzy" red slide with limited detail, or to go to extremes, projecting the white slide in a room with red background lighting?

Mr. W. N. Sproson (*in reply*): The N.T.S.C. colour television system makes (in principle) a separation of the luminance and chrominance components: when this is done, it is possible to severely restrict the bandwidth of the chrominance signal with almost no loss in the subjective sharpness of the picture. A Land colour picture with a red component of low bandwidth would not be satisfactory from the aspect of sharpness because the red channel carries both luminance and chrominance.

Messrs. Wilson and Brocklebank (*in reply*): No, we have not tried this variant of Land's projection. It seems to us that the range of subjects for which a two-colour system is acceptable is too narrow already to make still further reductions worth investigating. The extreme case of a white slide projected with an overall red light gives a result comparable to a black-and-white picture printed with red ink on pink paper, and would that be acceptable as a "colour picture"?

Mr. S. N. Watson: One of the most important factors, if not the most important, in considering the possible use of a new system for colour television is its effect on the cost of the receiver. There is no doubt that at the moment the cost of the colour television receiver is one of the principal causes of the relatively slow expansion of colour television in the United States. While my own opinion is that Land colour is most unlikely to provide an acceptable reproduction of the coloured scene for reasons, the effects of which were very ably demonstrated at the beginning of this discussion, it is nevertheless of great interest to know whether the cost of the receiver would be materially decreased by adopting this system instead of, say, a variant of the N.T.S.C. system. If I may add my own views on this matter, I should hazard a guess that so long as the display tube remains so large a fraction of the total cost of the receiver, any cheapening in the transmission system is likely to lead to only a small reduction in the cost of the receiver.

I should like to inquire whether the phenomena associated with Land colour are of a different nature when the pictures are moving, as would be the case should the system be used for colour television, as compared with the conditions in which still pictures are used, as has been the case in the demonstrations this evening.

Messrs. Wilson and Brocklebank (*in reply*): Adaptation phenomena hold equally well for moving as for still subjects, but the problem of moving from one scene to another with different lighting (e.g. from interior to exterior) would throw a greater strain on a two-colour than on a three-colour system. The Land system as it is cannot handle outdoor scenes, even still ones.

Mr. Sproson (*in reply*): The demonstrations given this evening have used static pictures: on the basis of an experiment I witnessed at the Colour Telecine Department of the B.B.C. Lime Grove Studios, I have no hesitation in saying that the colours generated by moving pictures in "Land" colour show no significant difference to those we have seen tonight.

Dr. D. Maurice: I am interested in Dr. Land's finding that many different colours can be seen on the basis of his system if two primaries, one each side of yellow, are used. That there is something peculiar about yellow, apart from the obvious fact that it is roughly in the middle of the visible spectrum, is I think indicated by the fact that it does not look like the result of adding red and green, whereas, at least in my opinion, magenta looks as if it is half way between red and blue, and cyan most certainly looks as if it is half way between blue and green. I wonder if Mr. Sproson or Mr. Wilson are as impressed with this odd feature of the colour yellow as I am.

Mr. Sproson (*in reply*): I would agree that yellow has these interesting properties. Yellow also has the property of being nearly invariant (i.e. independent of adaptation conditions) in the sense that a pure wavelength of about 590 mμm is nearly always described as yellow.

The version of Land colour demonstrated this evening used white and red lights, but Land does claim that any two monochromatic wavelengths will suffice providing that a certain minimum separation is observed.

Messrs. Wilson and Brocklebank (*in reply*): This odd feature applies even more to the colour white, which certainly does not look like the result of adding red, green and blue. In general, additive mixture adds only the quantity of light, whereas the quality (colour) of the light does not add, but cancels out, so that the resulting mixture contains only what was common to the component colours. Yellow occupies the same unique position in the reduced gamut of a red-*plus*-green system, that white occupies in the full gamut of the red-*plus*-green-*plus*-blue system; the analogy of a fulcrum or point of balance in a system of forces is a helpful concept.

Mr. I. J. P. James (*Member*): I would like to ask Mr. Sproson whether on the chromaticity diagram the elliptical curve enclosing "similar and identical" colour names applied mainly to male or female observers. Since female observers are much more discriminating, it is presumed that the area for females would be much less than for males. A method of colour reproduction which does not treat the majority of viewers equally is surely not to be recommended.

It seems to me that members of panels associated with decisions on colour television broadcasting should be examined for colour blindness.

Mr. Sproson (*in reply*): Our tests used both male and female observers and we found, as Mr. James comments, that the women are more critical (see Table 4). Figure 6 is based on the results of Tables 1 and 4 so that it gives more weighting to the observations of the men. All our observers were tested for colour blindness and I agree with Mr. James on this point.

Mr. C. Ridgers (*Associate Member*): What are the reactions of a colour-blind person when viewing a two-colour Land projection, and are those reactions different to those experienced when they view a three-colour picture?

It has been suggested that the Land process could be adapted to television and give a cheaper and easier maintained system than the usual three-colour one. A further speaker then said that the system would be inaccurate and that the public would not like it as a consequence.

A three-colour system is only as accurate as the coloured filters used, and the setting of the various controls. Also, since the public has quite happily accepted high-fidelity sound with a top cut at 4 kc/s, "soot-and-white-wash" television pictures, and 202½ line pictures, they would be, no doubt, quite happy with Land two-colour television.

Messrs. Wilson and Brocklebank (*in reply*): No actual record of a colour-defective's reaction to a Land projection is known to us, but it could be predicted. A tritanopic dichromat would see it as identical to a three-primary picture (i.e. in "full colour"), but a protanopic or deuteranopic dichromat would see it as identical to a monochrome picture (i.e. in "black-and-white"). Since all the chromaticities present in the Land projection fall in a segment of line joining the points representing the two primaries, these will all be perceived with the same colour if this line coincides with the colour-defective's confusion locus. Thus a two-colour system using blue and yellow primaries would appear fully coloured to a protanope or deuteranope, but black-and-white to a tritanope.

The experience of the photographic and cinema industries has been that the public holds strangely unpredictable but very strongly held views about colour, so that one cannot be sure how happy they would be with a two-colour system.

Mr. Sproson (*in reply*): The answer to the first question depends, I suggest, on the type of colour blindness. From the confusion loci corresponding to the different types of colour blindness, I should expect the protanope to lose nearly all colour sensation from a "Land" colour reproduction: deuteranopes and tritanopes would suffer some reduction in the range of colours they normally experience.

I accept the statements that reproduction in black and white television and in sound radio are often gross distortions of the original, but I doubt whether a system with the limitations of "Land" colour would be considered satisfactory. Monochrome television can give a good range of tonal values—it is not bound to be "soot-and-whitewash": Land colour has a restricted range of colours even when it is functioning at its best.

Mr. K. Hacking: The results of the various experiments with two-colour systems such as Land colour seem to indicate that the loci of observed colours is a surface located within the total colour solid normally used for the objective representation of a three-colour system. Do the authors think that it would be possible to carry out colour matching experiments in order to locate more precisely such a surface which would include both the subjective and objective realms of the system?

The same physiological processes in human vision, such as simultaneous contrast, which must account in part for the "extended" colour gamut observed in two-colour systems, must also be operative when viewing a three-colour system; but with a three-colour system we are not easily able to devise a co-ordinate system in order to represent the subjective effects. It seems to me that with any given two-colour system we have, theoretically at least, the opportunity to map out the subjective colour realm which it initiates within the three co-ordinate framework already devised for objective three-colour representation.

Messrs. Wilson and Brocklebank (*in reply*): The surface representing all colour points produced by a two-primary system is very easily obtained in an objective colour space (e.g. in the C.I.E. system); it is only necessary to know the chromaticities of the two primaries. But this does not give information about the subjective colour perceptions, any more than a C.I.E. specification does. The difficulties of relating perceptions to specifications based on colour-mixture data are no simpler for a two-primary system; indeed it cannot be done at all except in a three-co-ordinate framework.

Mr. Sproson (*in reply*): Certainly the various subjective effects such as simultaneous contrasts are present in viewing three-colour as well as two-colour images. I would agree that a suitable reference scale is provided by the three-colour solid to include all the colours generated by "Land" two-colour. Matching experiments could locate the two-colour surface within the three-colour reference solid and may be a binocular matching technique such as used by Hunt would be helpful in making the measurements.

Mr. B. W. Osborne (*Associate Member*): Though it may be considered that the Land presentation is not acceptable for accurate colour rendition, in place of a full three-primary display, is there not, nevertheless a possible case for its application if advantageous in low-cost colour receivers. This would be subject to the proviso that the transmitted signal is still capable (as with any likely adaptation of the N.T.S.C. system to present or future British standards) of providing us with the necessary information for the use of three-colour primaries on more elaborate receivers.

The cost of introducing colour lies more in the receivers than in the transmitters; and the cost of the colour receiver lies more in the actual display than in the associated circuitry. The "Land" receiver has a simpler problem of image registration (there being only two pictures to superimpose), and, as has been pointed out, would not be so liable to suffer from the danger of green or blue flesh tones. It might at least appeal to impecunious males.

Mr. Sproson (*in reply*): A television receiver operating on "Land Colour" would be somewhat simpler to manufacture: if the transmission were in full-colour, however, the colour information would effectively have to be processed so that R, G and (B) signals were obtained, i.e. synchronous demodulation would be essential. The display tube is the most expensive part of a three-colour receiver and a two-tube receiver (red and white) with a semi-reflecting mirror is a possibility although a somewhat cumbersome arrangement. In view of the relatively poor colour rendering of two-colour, it is very doubtful whether any manufacturer would be prepared to "tool up" for production of two-colour shadow mask, or beam-index tubes. This tube would suffer from uneven usage of its two guns for any monochrome transmissions, since the "red" gun would not be required.

U.D.C. 621.396.933

Automatic Techniques in Civil Air Line Communication Systems

Presented at a meeting of the Radar and Navigational Aids Group, held in London on 4th January 1961.

By

W. E. BRUNT†

Summary: The applications of automatic data transfer techniques to the airborne components of the aeromobile service are concerned with "machineable" methods of modulation and coding, and with input/output devices of a much higher degree of sophistication than such familiar transducers as microphones and telephones. The internal characteristics of the traffic which are important in the selection of coding methods can be seen from a system of classification. Aeromobile traffic can be divided into three categories, whose internal characteristics are so different that each demands its own particular coding technique. The optimum solution appears to demand the use of three sub-systems, each properly related to the others so that the end result is a single system. One of these sub-systems already exists, namely the radio telephone; the others so far exist only in elementary or experimental forms. The methods of integration, both procedural and functional, which are required to ensure that the three sub-systems are assembled into a single coherent system are discussed.

1. Introduction

Since the title of this paper may be subject to a variety of interpretations, it will be as well to define it as it is to be discussed.

Firstly, the scope of the discussion is limited to aeromobile communications, with particular reference to aircraft equipment.

Secondly, whilst "automatic techniques" are employed in all modern aircraft communications equipment, the paper is specifically concerned with what is now colloquially known as "automatic data transfer".

In short, by "automatic techniques" we mean "machinable" methods of modulation and coding, by which, for example, machine may talk to machine, as opposed to the present systems which inherently demand that man shall talk to man. We also mean methods which will permit the much more efficient transfer of information from man to man, i.e. the somewhat complex field of input, output and translation devices.

Having now defined the subject, it will be as well also to define the terms in which it is to be discussed.

Firstly, there would seem to be little point in discussing any system without first having established the ambience in which it will be required to operate. In the present case, this means that we must first state the nature of the total communications requirement in the aeromobile service, in terms that are meaningful in the present context. This will be attempted herein, even if, of necessity, in a rather superficial manner.

Secondly, the contents of this paper are the responsibility of the author; in no sense can they be taken as representative of agreed requirements, or even thinking, of the airline industry. For, in fact, whilst there is rapidly increasing support for automatic communications in principle, there is as yet no agreed statement of the requirement. And in view of the co-operative nature of aeromobile communications, and the multiplicity of state authorities and airlines involved, such agreement is a pre-requisite to real progress.

2. The Nature of the Requirement

Aeromobile communications are, of course, a particular form of mobile communications in which the vehicles involved are aircraft. They are thus, by definition, radio communications.

In the great majority of cases, one terminal will be an aircraft (not necessarily airborne) and the other will be a specially-assigned ground station. There is, however, a small amount of traffic from aircraft-to-aircraft or ground-to-ground. With some small exceptions, this traffic is now conducted by amplitude-modulated radio telephony.

In the main, although again there are some exceptions, these communications are conducted within the h.f. band 2·0–20 Mc/s and the v.h.f. band 118–

† Communications Superintendent (Systems), British Overseas Airways Corporation, London Airport (Heathrow), Middlesex.

136 Mc/s. V.h.f., because of its inherently greater efficiency and reliability, is the preferred mode wherever practicable, but for the major international routes the longer range h.f. mode is essential.

The motivations behind the desire to "automate" these communications may be complex in detail, but in essence they are identical with those which arise for instance in the manufacturing industries, namely the need to improve productivity, control and quality. Given an open mind, one can look at the present system from almost any point of view and yet come to the same conclusion; as the sole means of communication, r.t. leaves very much to be desired.

In order to provide a satisfactory basis for the choice of suitable alternative methods, it is necessary to reveal the relevant characteristics of aeromobile communications. To this end, the author's organization adopted the following procedure which, whilst relatively simple, has done much in providing a better understanding of the problem.

Firstly, aeromobile communications can be shown to consist of three general types or categories:
(a) Advisory—one way, ground-to-air
(b) Routine—two way, air-ground-air
(c) Non-routine—two way, air-ground-air.

Category (a), as its title implies, is in the nature of advice to those who wish to avail themselves of it. It is largely broadcast, as opposed to being addressed, and since it contains no element of command, there is no requirement for specific response. A particular example of category (a) is the meteorological broadcast.

Category (b) is quite different; before any flight begins it is known that these messages will have to be exchanged, their form is known and their content and time of transmission will be known within limits. A typical example is the Air Traffic Control Position Report.

Category (c), on the other hand, consists largely of unpredictable traffic, i.e. it arises as a result of events which were not planned into the operation. It covers an extreme range of possibilities, and thus makes particular demands for flexibility. It is hardly appropriate, in this case, to talk of "typical" cases, but it is only necessary to dwell for a while on the range of information which might have to be handled under the general heading of "Emergency" to see the implications.

We can now take the next step in classification, one which exposes some most informative contrasts and which begins to illuminate such aspects as coding and input/output devices.

Broadly speaking, the vertical or procedural classification can be expanded by horizontal classification in functional terms, as shown in Table 1. Clearly, it cannot be claimed that this is a rigorous classification; nevertheless, we have not so far discovered that it has any major weakness.

Table 1

Category	*Direction*	*Volume*	*Urgency or "Priority"*	*Vocabulary*	*Predictability*
1. Advisory	G/A	High	Low	Medium	Medium
2. Routine	A/G/A	Low	Medium	Small	High
3. Non-routine	A/G/A	Low	High	Large	Low

The most revealing thing to emerge from this classification is the lack of similarity, in the three categories, in terms of their internal characteristics. Without going further into the matter at this point, suffice it to say that this study has lead to the conclusion that no one "method" of communication can hope to satisfy the requirement, and that the ideal solution will almost certainly employ three sub-systems, each based upon a particular method or technique. Nevertheless, we also believe that there must be a large measure of integration, both electronic and procedural, so that the sub-systems are truly co-operative, each playing its part in optimum fashion. Only in this way can we hope to develop the ideal of a single communications system.

The remainder of this paper will be devoted to discussion of a system concept based upon the foregoing considerations.

3. Choice of Method

Implicit in the discussion so far is the inference that each of the three categories of traffic demands its own and quite individual communication technique. It would at first sight seem logical to discuss them in the order in which they have been stated. But this would be to risk an anti-climax, for, in the present context at least, there is so little to be said about the system required for category (c). For we already have it.

In spite of the known deficiencies of radio telephony, it has certain very real advantages. From the point of view of the user, there is no coding problem, it will handle an almost infinite vocabulary, is easily implemented and widely available. Its psychological value is high, and for this reason alone it will be retained. This being so, we might as well make use of it to reduce the severity of the problems which would otherwise arise in the fields of coding and of input/output devices. In other words, by saying that all that cannot be conveniently handled by sub-systems (a) or (b) will be handled by radio telephony.

In short, whilst everything practicable must be done to minimize the use of r.t., it will nevertheless continue to play an important role, and for this reason the search for improvement must continue.

In approaching the system requirements posed by category (a), B.O.A.C. and its partners in the enterprise laid down certain basic requirements, namely:

(1) The airborne equipment should demand only minimum attention and skill in operation.
(2) The information produced should be self-evident.
(3) It should produce a permanent record in convenient form.
(4) First priority should be the long-haul application.
(5) The solution would be limited to broadcast information in the first instance.

Furthermore, it was agreed that since meteorological information accounted for such a high proportion of category (a) traffic, attention would be concentrated on that aspect for the time being.

As the result of an extensive study, it was concluded that of those likely to be available in the immediate future, the teleprinter was the most practical answer. To state only some of the reasons, since operational simplicity was essential and reliable ranges of the order of 1200 miles were implicit, then low-frequency propagation was mandatory. Thus, available bandwidths were very limited. This, coupled with an apparent requirement for print-out at a rate of not less than 60 words/min, was sufficient to rule out some techniques, e.g. facsimile.

Furthermore, the teleprinter is already a principal means of passing meteorological information over the appropriate ground networks, and thus the element of compatibility promised not only to speed dissemination but also to offer important economies to the authorities.

But perhaps the determining factor was that in addition the teleprinter offers certain important degrees of flexibility. It offers a very wide vocabulary, and the "tightness" of coding can be varied over a significant range. This makes it possible to optimize coding in relation to the quality of available channels, the quantity of traffic to be passed and the necessary degree of integrity. The following examples illustrate the point.

(1) EGLL AERO 2250 140/13 8 NM 4/8 6000 8/8 9000 QMU 1412.

(2) LONDON AERO TWO TWO FIVE ZERO ONE FOUR ZERO DEGREES ONE THREE KNOTS EIGHT MILES WEATHER NIL FOUR OCTAS SIX ZERO ZERO ZERO FEET EIGHT OCTAS NINE ZERO ZERO ZERO TEMPERATURE ONE FOUR DEW POINT ONE TWO.

(3) LONDON AIRPORT WEATHER AT TWENTY TWO FIFTY HOURS. WIND DIRECTION ONE HUNDRED AND FORTY DEGREES, WIND SPEED THIRTEEN KNOTS, VISIBILITY EIGHT MILES, NO RAIN. CLOUD BASE FOUR EIGHTHS AT SIX THOUSAND FEET, UNBROKEN COVER AT NINE THOUSAND FEET. TEMPERATURE FOURTEEN DEGREES CENTIGRADE, AND DEW POINT TWELVE DEGREES CENTIGRADE.

The first example is a standard meteorological broadcast, in which apart from the degree of compression two features are notable, i.e. the high proportion of numerals and the very low degree of contextual protection. An error in this message has a good probability of going undetected.

The second is the same message expressed in another way, the effect being to reduce almost to zero the probability of undetected error.

In the third example, the same message has been further expanded into completely colloquial form. Protection is almost complete, but the effective information capacity of the circuit has been reduced by a factor of about 6.

It is not difficult to see that in the first case an error-rate of 1 in 1000 might be unacceptable, in the third case an error-rate of 1 in 20 could be tolerable.

A system meeting the above-stated requirements has now been in operation for some years, albeit on an experimental basis. As a result of very extensive use (and B.O.A.C. now have 25 aircraft equipped), it has been established that

(1) Reliable coverage of the whole of the North Atlantic can be provided from two stations, one on either side.
(2) Only one frequency per station is required for 24 hour service throughout the year.
(3) When h.f. communication is disturbed because of poor ionospheric conditions, the reliable range of the l.f. radio teleprinter (r.t.t.y) system is appreciably extended.

The system characteristics, which are very simple, are as follows:

Modulation	frequency shift keying
Shift	60 c/s ($\pm$ 30 c/s on nominal frequency)
Speed	60 words/min
Receiver bandwidth	110 c/s to 6 dB
Carrier	110–150 kc/s
Signalling code	5/7½ unit Baudot.

The printer is a modern but nevertheless conventional example of its kind. The receiver, specially designed and produced for the project, is remarkable more as an engineering exercise than for the system principles upon which it is based.

Since bandwidth and shift have both been reduced to the minimum in order to gain maximum protection against noise, a high degree of frequency accuracy and stability is essential. The requirement is that drift and initial setting error combined should not exceed 4 c/s, and this is generally met. In the latest versions, receivers are completely transistorized and further protection against noise has been obtained by employing an antenna consisting of critically-coupled crossed loops.

But although a good deal has been done, this system is still relatively simple, and satisfies only a part of the requirement. Is there scope for further development, and is there a need for it? So far as B.O.A.C. is concerned the answer to both questions is an emphatic "yes". The most attractive lines of development appear to be as follows:

(1) The printer, good as it is, is bulky. It is the only item of the system to pose an installation problem. A much more compact printer would result in greatly increased support for the system.

(2) The system should provide for selective routeing as well as broadcast. We need selective addressing.

(3) There is a case for short-range use on v.h.f. In this case the stringent modulation conditions appropriate to l.f. do not apply, and it should be possible to use the existing v.h.f. receivers.

Given success on all three of the above, it would appear that there will be very little of category (a) traffic which the system could not handle, and it will provide valuable support in the operation of the complementary sub-systems.

By and large, category (b) traffic is concerned with air traffic control, and the range of information to be handled is quite small and consists of two parts:

(a) Factual information, i.e. three-dimensional position reports, and

(b) Executive information, i.e. instructions and requests for information or further instructions.

The first of these will, for preference, be performed on a machine-to-machine basis, i.e. the aircraft *equipment*, rather than crew, will report position to the automatic data processing equipment in the a.t.c. centre. The second is essentially a matter of man-to-man communication. There is thus a requirement for two distinct classes of input/output devices.

In considering the informational characteristics of this sub-system, the first requirement to be noted is that of station identity, or "address". Ways can be seen by which the ground station can be relieved of this requirement (e.g. address may be implicit by virtue of some other characteristic of the transmission) but it seems that all messages must contain the identity of the concerned aircraft.

So we have to select a suitable system of identification. Procedurally, there is already a choice of methods; all aircraft carry a national registration, expressed in a total of five alphabetical or alphanumeric characters, and all civil flights are identified by a flight number, which generally requires five alphanumeric characters but sometimes six. Many other methods are possible, of course, but there is great advantage to be derived from the use of that form which is already part of the a.t.c. system, i.e. the flight number. Thus, some 25–30 bits are required for identity. It is likely that this will be the longest component in this class of message, and may sometimes amount to more than 50% of the total.

Coming now to the three-dimensional position, altitude presents no great problem. In all probability, we will continue to express altitude in feet; if a change is made, it will be to the metric system. Similarly, the minimum increment used for a.t.c. purposes is 500 feet. So, making all due allowances for the altitudes at which the supersonic aircraft of the future may operate, altitude requires no more than 10 bits for adequate resolution.

The lateral parameters pose a more difficult problem with two major components. In the first place, there is no fixed increment, as in the case of altitude. We must choose increments based upon the precision of position-determination, and this may vary from as little as 0·5 nautical mile in high-density areas to as much as 20 nautical miles over the oceans and deserts. Secondly, the systems of co-ordinates and procedures now in use do not appear to be ideal. The language of the system used in high-density areas, i.e. the use of nominated check-points ("over Epsom range", "over Benbecula beacon") is not a logical one and thus is not directly amenable to the use of machineable codes. That used in long-range over-ocean flights, i.e. geographical coordinates, is completely logical, completely unambiguous and, consequently, highly redundant. To express a lateral position in these terms and with adequate resolution would often require about 36 bits, whereas the true informational content of such a fix would seldom exceed about 9 bits.

If a new procedure is to be considered, then there is good reason to take advantage of the fact that a detailed flight plan is prepared in advance of every flight. This is, in effect, a minute-by-minute prediction of the position of the aircraft, subject to certain unknown variables. It is thus at least theoretically possible to report position as an error from flight plan. This appears to be a logical and economical approach, since it is completely systematic and would rarely require more than 10 bits.

These considerations cover the main function of automatic position reporting; there will probably be some requirement for further elements of information, such as those required to reveal the true meteorological ambience, but these are unlikely to pose any real problem.

So much for machine-to-machine functions. In the man-to-man case it can be said that either pilot or controller must have the ability to pass an instruction or advice, to signify acceptance or rejection of such, and to initiate a request for advice. These instructions or advices may relate to altitude, bearing, airway, holding pattern or speed, from which it follows that in each case there will be an appropriate dimension or identifier.

This describes the absolute minimum of facilities, of course; it may be desirable to increase the scope to include landing instructions and more complex control functions. But such additions must be examined with care, since a prime requirement is to ensure extreme operational simplicity; any such system will fail if it is easier for the pilot to pick up a microphone and speak.

There is much scope for ingenuity both in the selection of a suitable encoding scheme and in the design of input/output devices. On the former it seems probable that, in contrast to the category (a) case, we shall not transmit the actual message to be displayed, but rather the bare informational content. Proper provision for "addressing" will require that this be programmed prior to flight; the address must be automatically included without attention from the flight crew.

Finally, the equipment must be capable of universal application, irrespective of the frequency of propagation. If this involves, as it probably will, a change of bit-rate as between v.h.f. and h.f. modes, then this too must be automatic.

4. Integration

Having broadly described the three sub-systems, it now remains to show how they interact to form a coherent whole.

First it should be noted that system (b) is feasible only because system (a) is available to deal with high-volume traffic, and system (c) is available to deal with low-predictability traffic. It is also to be expected that system (a) may be able to make a substantial contribution to the performance of the other two where reliability is now most difficult to achieve, i.e. in the case of h.f. propagation. Plans are now far advanced to introduce the measurement of actual (as opposed to *predicted*) optimum working frequency for the North Atlantic h.f. aeromobile frequencies, and to broadcast the results on the r.t.t.y. channel at thirty minute intervals. Preliminary trials have indicated that this promises a most significant improvement in the quality and reliability of service. It may well foreshadow the day when the frequency-channelling of airborne equipment will actually be controlled from the ground.

These are examples of procedural integration, with some electronic overtones. But systems (b) and (c) also demand specific electronic integration; for not only must the coding and input/output devices be common to both h.f. and v.h.f. modes of propagation, but the carrier-generating and frequency-selection equipment must, of course, be common to systems (b) and (c). Figure 1 shows these interconnections.

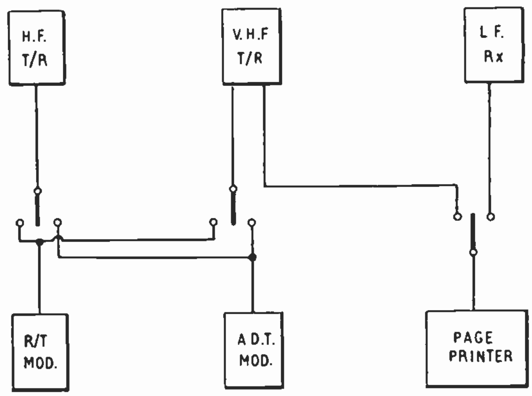


Fig. 1. Integration of aeromobile automatic communication systems.

The possibilities, of course, go far beyond what has been described. But perhaps enough has been said to show that the total complex is, by definition, a single system, since each of the parts is designed in relation to the capabilities of the others, and each has its own particular importance in contributing to an optimum solution of the total aeromobile requirement.

5. Conclusions

It has been shown that the nature of aeromobile traffic is such that no one coding technique is likely to suffice, that indeed a satisfactory solution is likely to require the application of three such techniques.

It has also been shown that of these three methods of coding, one can be stated with certainty, a second with some conviction, but the third can be discussed, at this time, only in broad generalities. This degree of uncertainty is likely to persist in the absence of an agreed international requirement, with particular reference to the future communication needs of air traffic control.

Manuscript received by the Institution on 14th December 1960.
(Paper No. 646.)

DISCUSSION

Under the Chairmanship of Mr. G. R. Scott-Farnie, C.B.E. (Member)

Mr. D. M. O'Hanlon (*Associate Member*): May I say straight away how pleased I am that Mr. Brunt's paper emphasized the significance of the "information" in the communications channel, with regard to automation, rather than methods by which automation may be achieved. Too often in the aviation business we are plagued by comparisons of excellent solutions to problems which either do not exist or have been inadequately stated.

In this connection, the relative proportions in Mr. Brunt's three categories affect the case. Could Mr. Brunt give any information on what the proportions are in the case of the long-haul operator?

For the short-haul operator in a high traffic density area with a large number of reporting points may be the case for automatic communications could be made on a reduction of workload for the pilot.

As a side issue, because some present may not be aware of it, would Mr. Brunt give some information on the "back-scatter" work being carried out on the North Atlantic routes to enable B.O.A.C. aircraft to be provided via the airborne teletype receiver with data on the best h.f. communication channel.

Mr. W. E. Brunt (*in reply*): In such studies as we have made, the proportions appear to be, very approximately, 80% in category (a), 15% (b) and 5% (c). If the proportions could be determined with precision, then it is probable that there would be differences as between one airline and another. Such differences, however, would not be such as to affect the choice of systems, except possibly in the case of the really short-haul operator, who might conceivably elect to omit sub-system (a).

With regard to the "back-scatter" sounding station at London Radio (Birdlip), the purpose of this application of back-scatter is, of course, to define the optimum operating frequency by *measurement* (as opposed to prediction). The arrangements now being made will permit soundings to be made on frequencies which are, for this purpose, substantially identical with those assigned to the North Atlantic aeromobile service, and to provide a visual display of skip-distance and effective working areas.

Preliminary tests have shown great promise, and suggest the possibility of a most significant increase in both reliability and quality of communication. If this promise is borne out, then the implications in respect of automatic signalling on h.f. are clearly most significant.

Mr. N. G. V. Anslow (*Associate Member*): I should like to point out the considerable saving in flight deck work load which can be achieved by having the reception of ground to air advisory information automated. An investigation was made in B.O.A.C. on our Trans-Atlantic Comet Mk. IV jet operations into the time spent on the flight deck dealing with h.f. *en route* radio communications, i.e. the passing of position reports and the reception of meteorological information etc. It was found that over 70% of the time spent on these communications had to be devoted to the reception of weather broadcasts. The fitting of a l.f. r.t.t.y. system can therefore materially contribute to reducing aircrew fatigue.

Mr. Brunt: Yes, and this agrees very closely with the results of our earlier study.

Mr. R. Travers (*Associate Member*): I have noted the large number of bits required for a position report from an aircraft and would like to ask whether the use of a modern continuous-wave hyperbolic navigational radio aid would perhaps reduce the number of bits required. For instance, could airborne position indicators be interrogated? I ask this question because with geographical co-ordinate position reporting the number of bits required will increase as the controlled vertical and horizontal separation required in high density areas decreases.

Mr. Brunt (*in reply*): The necessary number of bits is determined by two factors: (a) the resolution required, and (b) the degree of ambiguity which can be tolerated. No choice of system can reduce this minimum. In the example given in the paper, the geographical co-ordinate system shows up badly purely because it is completely unambiguous, each spot on the Earth having its own discrete identity.

Mr. G. W. Barnes: In contemplating fully automatic communication for the aeromobile service it is important to remember the part now played by the r.t. operators. At long ranges, using h.f. propagation, communication is established as a result of their judgement and time sharing on the same channel is possible because they listen before transmitting. In the automatic case it is expected that data will be transmitted as a result of the sequential interrogation of all the aircraft being controlled at any time. If it is assumed that h.f. propagation will limit the data rate to 50 bits per second and that the message content, including the identity, could be 100 bits, then allowing for switching periods an exchange of messages could take 5 seconds. A system capacity of 240 aircraft would mean a 20 minute interval between interrogation for any one aircraft.

Mr. Brunt mentioned the use of a loop aerial on the aircraft for reception of the long range m.f. teleprinter broadcasts. I should be glad if he would give more details of this and say how the performance compares with that of a wire aerial.

Mr. Brunt (*in reply*): In reply to the first question I would say that the average message is not likely to contain as many as 100 bits, and that the maximum number of aircraft to be served by any one h.f. channel is unlikely to approach 240; 50 is a much more probable maximum. For these reasons, a two-minute interval seems reasonable, with 5 minutes as an outside maximum.

Broadly, the particular loop aerial scores by virtue of greater protection against local electrostatic fields, so that whilst in the particular case the effective height was significantly less than that of the wire, this was more than balanced by the improvement in signal/noise ratio. The overall improvement has not been precisely determined, but appears to be of the order of 6–10 dB.

The aerial consists of two loops with axes crossed. They are accurately tuned and balanced, and are critically coupled so as to produce a substantially circular polar diagram. The loops are wound on ferrite bars, and the whole assembly is "potted" to produce a "pancake" of about $20\frac{3}{8}$ in $\times$ 11 in $\times$ $1\frac{5}{32}$ in, which may be mounted directly on the aircraft skin.

Mr. M. Settelen: In view of Mr. Brunt's very considerable experience serving on international committees, is he able to indicate how he thinks international agreement could be achieved on the whole matter of automatic data transfer, particularly in view of the fact that it will affect not only the airborne equipment dealt with by the airlines themselves, but also the air traffic control authorities in many countries.

During Mr. Brunt's paper he specifically referred to the use of high frequencies for long distance communication. How does Mr. Brunt reconcile this with a recent decision taken by another major world airline that all h.f. communication airborne equipment would be removed from their aircraft by the end of 1962? This airline has, during the past several years, undertaken considerable research in the use of long distance v.h.f. and apparently are obtaining very impressive results.

Mr. Brunt (*in reply*)*:* This is most difficult; however, I believe two things are necessary. First, the air traffic control people must establish, through I.C.A.O., an agreed requirement for aeromobile communications as a function of a.t.c. with particular reference to automation. Secondly, I.C.A.O. must produce an appreciation of the integrated requirements for the whole of aeromobile communications. Until these two steps have been taken, it is difficult to see clear prospects of *orderly* progress.

With regard to v.h.f. replacing h.f. entirely, I am not aware of any such decision by an *airline*, although individuals may have made optimistic statements. I would only comment that such a development would in no way adversely affect the general scheme which I have presented; it would simplify the equipment and reduce some of the more difficult problems. I would add that the high reputation that v.h.f. now enjoys springs largely from the fact that its operation is quasi-optical, and care must be taken not to introduce those very uncertainties which beset h.f.

Mr. R. D. Jones: I am directly concerned with the equipment that will have to be installed in aircraft to provide the automatic communication facilities that Mr. Brunt has described. It is obvious that this equipment will add to the complexity of radio installations and therefore to the cost of maintenance, so I should like to ask Mr. Brunt what advantages to the airline he can foresee to justify the extra cost of carrying the equipment. Reduction of pilot work load has been mentioned but this is not necessarily of direct benefit to the airline. It is possible that operational efficiency will be increased, by reducing flight time for example, or it may be that automation will become essential with higher aircraft speeds or with greater traffic density in terminal areas.

Mr. Brunt (*in reply*)*:* There is no single justification; it is a composite which contains all those possibilities mentioned by Mr. Jones, and many more. To mention three specific benefits, we aim to reduce, if not eliminate, the need for communications expertise in the flight deck, to provide a much more adequate flow of information for air traffic control and other purposes, and greatly to reduce channel-changing. The result of these three things alone will be more economical control of flight whilst freeing the crew from "non-flying" duties. In particular, I believe that measures of this kind will be quite essential to the supersonic aircraft.

Air Commodore C. A. Bell: Does Mr. Brunt consider that a further economic advantage of the system of communications which he has described would be a better ordered and better disciplined use of the spectrum, with consequent advantages to all users of radio?

Mr. Brunt (*in reply*)*:* Most emphatically, yes. So far as v.h.f. is concerned, most of the channel-changing which now goes on can be ascribed to the incredibly low efficiency of channel utilization. Something like 7 or 8 channel-changes are necessary on the approach to New York, with the consequence that much more time is spent in setting up channels than in using them. And yet, given a new system having an efficiency of no more than, say 5% of theoretical, a single frequency could serve the entire approach and landing. Clearly, the benefits must also be felt by other users of radio.

Mr. B. J. Infield: One of the most urgent problems with which we are now faced is that of collision avoidance. It is also one of the most difficult to solve and all the work on the subject to date leads to the conclusion that a satisfactory airborne solution is unlikely to be forthcoming in the forseeable future. If, as seems probable, collisions will have to be prevented by control from the ground, does Mr. Brunt think that automatic position reporting can play a useful part?

Mr. Brunt (*in reply*)*:* If I am correct, then a properly designed automatic communications system will provide more accurate data and a higher data renewal rate than present systems. This, in turn, should provide air traffic control with a more accurate record of position, both absolute and relative. Therefore, the probability of collision must in some degree be reduced, and my answer must be yes, it can play a very useful part.

Mr. G. R. Scott-Farnie (*Chairman*) thanked Mr. Brunt on behalf of the Radar Group for his most interesting and stimulating paper. He felt that it was very doubtful whether the development of automatic techniques for airline communications would be applicable to the present generation of aircraft largely due to two basic problems: (a) the agreement necessary between the airlines themselves as to what was required and, (b) the international agreement necessary between states for implementation.

Airborne radio teletype for meteorological information—a relatively simple form of automatic communication—had taken nearly ten years to implement, was confined to relatively few airlines and was only in operation on the North Atlantic route.

For the next generation of aircraft however—the supersonic—automatic techniques for communications would be a "must" and to meet this requirement it is essential for airlines and states to establish their requirements immediately.

Developments in Aviation Electronics

Electronic Equipment in Aircraft

During the past eight years, aviation electronics has earned Great Britain £23¼ million in export sales, a figure which does not include the value of equipment installed in aircraft for export. This was stated recently by the Society of British Aircraft Constructors.

The demand for electronic equipment increases constantly. Military and civil users seek more and more assistance in performing special tasks, not only in navigation, communications and weapons control, but in aircraft systems of all kinds. For example, a medium bomber built in 1939 carried radio and electronic equipment weighing less than 150 lb: the corresponding figure for a V-bomber of to-day is about 10,000 lb or nearly five tons. A wartime fighter carried two miles of wiring and its radio equipment included twenty valves: a present-day fighter has three or four times as much wiring and 150 radio valves, while the V-bomber has twenty miles of wiring and more than 1000 valves.

The need to reduce equipment weight in commercial aircraft is thus of particular importance, for every extra pound of weight costs money. It has been estimated that each additional pound in a trans-atlantic airliner costs the operator £1000 a year. Great Britain, the country which invented and developed radar and which has been responsible for many modern aviation electronics developments, is contributing many new developments in satisfying this ever-growing demand for the minimum weight in airborne radio equipment.

Clearly, miniaturization is of the utmost importance. Playing a key role in the campaign against surplus weight are the transistor and other semi-conductor devices. Apart from their very small size the reduction in power generating equipment and avoiding the need for forced ventilation by power-driven blowers are factors which represent a considerable saving of weight and space. Examples of what can be achieved is an airborne weather radar, introduced recently by Ekco Electronics, which has a range of 150 miles ahead of the aircraft, yet which weighs only 57 lb. Another is Marconi's 60 Series receiver which is forty times lighter than equivalent equipment using valves.

Increasing Application of Secondary Radar

Another interesting development in aeronautical electronics is the recent announcement that British Overseas Airways, Air France and Air India International are to install Cossor secondary radar transponders in their Boeing 707 jets. The transponders are of course the airborne part of the Secondary Radar System whereby aircraft fitted with them can identify themselves to the Air Traffic Controller on the ground in a positive manner and without any distraction to the pilot.

The principles of the Cossor secondary radar system were described in the *Journal* some years ago.† It will be recalled that a transmission from the ground equipment "interrogates" the airborne transponder which then replies with a train of pulses. Sixty-four different pulse combinations are available. The ground equipment resolves the pulse combination and presents it to the Air Traffic Controller in a form which identifies the aircraft. At the same time the usual p.p.c. presentation provides range and bearing information for the same aircraft.

In the U.S. the carrying of secondary radar transponders is already compulsory for civil aircraft flying altitudes above 25 000 ft. Similar arrangements are under consideration for the U.K. and Western Europe. A secondary radar ground equipment, also supplied by Cossor, is now being evaluated at London Airport by the Ministry of Aviation.

Pilot's Eye-level Electronic Presentation (PEEP)

A "head-up" type of navigation display which is projected on to the wind-screen of an aircraft has been engineered by Rank Cintel Ltd. to meet the requirements of the Ministry of Aviation, for a collimated display of information from cockpit instruments. The Cintel system, known as Pilot's Eye-level Electronic Presentation (PEEP), is derived from a head-up type of display developed by the Royal Aircraft Establishment. The equipment comprises a pilot's display unit and a waveform generator. In its simplest form PEEP serves as a flight indicator whereby the pilot sees a collimated display of aircraft altitude, height, etc.

Telephone Calls from Aircraft

The existing United Kingdom h.f. long-range radio telephone service from ships at sea has now been extended to telephone calls from aircraft. The service will be available to any airline or privately owned aircraft, subject to satisfactory tests. The charges for calls will depend upon the position of the aircraft.

In the interests of safety any airline wishing to take part is asked to give an undertaking that an uninterrupted effective watch on the air traffic control channel will be kept at the same time as the telephone calls are made. Some European countries and an American telephone company operate similar services.

† K. E. Harris, "Some problems of secondary surveillance radar systems", *J. Brit.I.R.E.*, **16**, pp. 355–82, July 1956.

Miniature Medium-duty Sealed High-quality Relays

By

N. E. HYDE†

Presented at the Symposium on New Components held in London on 26th–27th October 1960.

Summary: In recent years relays have become smaller and lighter in weight and in the majority of cases they are sealed. Unfortunately, reliability has not been improved, and in many cases it became worse. This paper deals with small, sealed, high performance relays designed for use in reliable electronic equipment.

1. Introduction

The use to which these relays will be put introduces a specification which includes proper contact functioning during Service conditions involving operation over a voltage range of $\pm 20\%$; temperature range of $-55°$ to $+125°$ C; relative air density 0·374 to 0·003 (equivalent to an altitude at 130 000 ft); 56 days at 40° C with relative humidity (r.h.) 95%; six cycles at 25° C rising to 55° C with r.h. at 80%; and finally shocks and centrifugal accelerations of 100 *g* and vibration up to 100 *g* over the frequency spectrum 30 to 5000 c/s. Sealing tests, shelf life and endurance tests are also included.[1, 2]

It is the author's opinion that earth-bound equipments should not tolerate less reliable components than airborne equipments. It is assumed that human lives may depend upon the proper functioning of both types of equipment and therefore reliability of the components must be maximal in both cases. To accomplish this, so far as relays are concerned, automatic monitoring of all contact functions during endurance testing must be carried out.

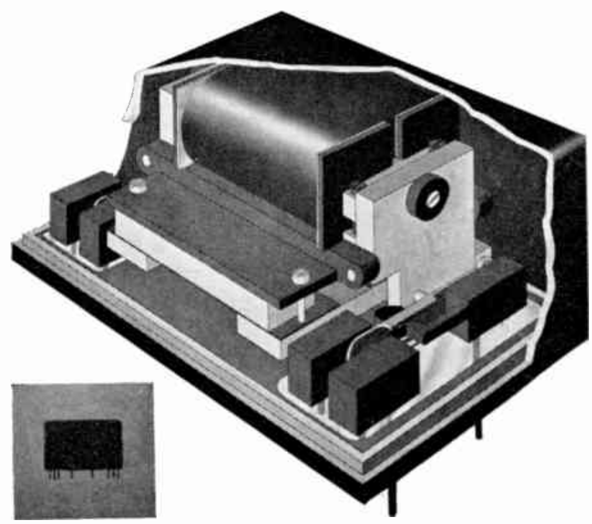
The criterion for determining the proper functioning of a relay will be that its contact resistance shall not exceed 100 milliohms when measured at a maximum voltage of 10 millivolts at 200 cycles per second. To achieve such a low resistance, it is preferable that the relay contact should be sealed within a "separate" compartment as in double-compartment relays.

Sometimes the most onerous modern requirement is a limited space and severe vibration, especially in aircraft and missiles, where automatic guidance and throttle control are vital. For this reason the paper is limited to relays smaller than any in the Services "Blue Book",[3] and to relays capable of withstanding high levels of *g* over a frequency band of about 30–2000 c/s.

† Ministry of Aviation, Royal Radar Establishment, Great Malvern, Worcestershire.

2. Century G "Relay"

The relay shown in Fig. 1 was probably the first two-changeover d.p.d.t. relay designed to withstand 100 *g* acceleration in this frequency band. Its size, weight, operating characteristics and other information are given in Table 1.



1·109″ × 0·726″ × 0·7″

Fig. 1. The "Century G" double-pole double-throw miniature relay. (Developed by New Electronic Products Ltd. on behalf of the Ministry of Aviation).

This relay is seen to comprise two well-balanced armatures suspended about their centre-of-gravity by large pinion bearings. Fixed rigidly at right angles at one end of each armature is a novel linkage system which combats linear and torsional vibration by allowing the armature to rotate only in opposite directions and at the same instant.

One end of each armature operates a restrained spring changeover contact assembly. Each end of each armature operates contacts in the 4-changeover variety of this relay. A permanent magnet aids the

Table 1

Requirements	Smiths PCR	Fortiphone G100	Others (Average)
Size: 1 changeover	0·5″ × 0·25″ dia.	—	—
2 changeovers	—	0·9″ × 0·8″ × 0·375″	1″ × 0·8″ × 0·35″
4 changeovers	—	0·9″ × 0·8″ × 0·75″	—
Operate Power:			
(nominal)	350 mW	125 mW	1000 mW
(minimum)	300 mW	60 mW	300 mW
Operate Time	2 ms	5 ms	5 ms
Release Time	2 ms	8 ms	5 ms
Bounce (milliseconds)	<1 ms	<2 ms	2–100 ms
Contact Resistance (milliohms)	70 mΩ (max.)	70 mΩ (max.)	25–200 mΩ unstable
Contact Rating (a)	1 amp, 30V	1 amp, 30V	1 amp, 30V
D.C. (b)	Dry circuit	Dry circuit	Dry circuit (unstable)
Life	(a) 10^6 operations (b) 10×10^6 operations	(a) 10^6 operations (b) 10×10^6 operations	(a) 5×10^5 operations (b) — operations
Vibration (30–2000 c/s)	50 *g*	50 *g*	20 *g*
Temperature	− 55° to + 125° C	− 55° to + 125° C	− 55° to + 125° C
Humidity (RCS.11)	H1	H1	H1
Sealing	Double compartment	Double compartment	Single compartment

return of the armatures and holds them securely in the un-energized condition. The permanent magnet lies parallel to the electromagnet, and this presents some interesting and incidental advantages.

The differential fields are easily adjusted by movement of the magnet or by fitting various strength magnets. This facility, coupled with reversal of the electro-magnetic field, changes the operate and release times considerably so providing the facility of "two-relays-in-one". Contact operate and release characteristics remain fairly constant and are practically independent of the operate and release speed of the relay. This is shown in Fig. 2.

3. "Postage Stamp" Relay

A common form of relay construction has been called "postage stamp" by virtue of its proportions and dimensions of the larger surface. The relay, which is shown in Fig. 3, is a direct replacement for the many relays of American origin which investigations by the author have shown to consume considerably more power.

This relay has two coils which are wound on the arms of a single piece "U" shaped electromagnet between the poles of which is pivoted a dynamically-balanced armature. Great care has been taken to suspend the armature so as to prevent end and side play in the bearings.

Optimum pole, core and armature dimensions have been incorporated in order to achieve the mentioned sensitivity in the available space. An enlargement at the closed end of the yoke increases the pull considerably and also acts as a heat sink.

The contacts incorporate the constrained spring type mechanism which successfully withstood very

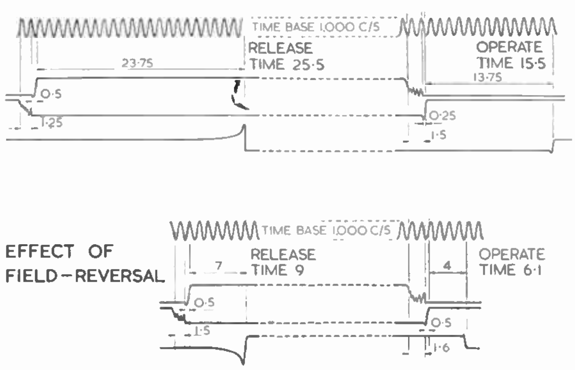


Fig. 2. Operate and release characteristics of the "Century G" relay.

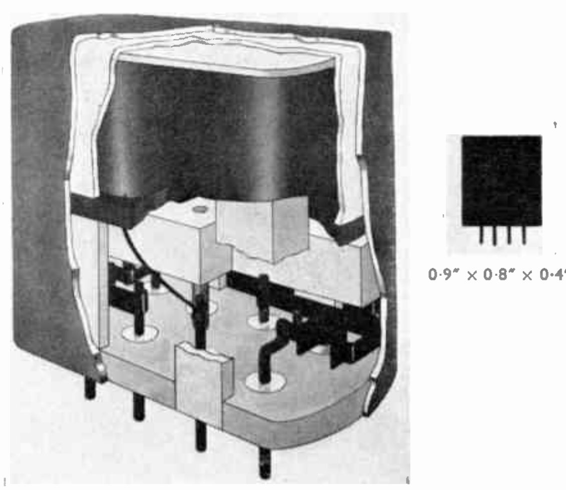


Fig. 3. The "postage-stamp" relay. (Developed by Fortiphone Ltd. for the Ministry of Aviation.)

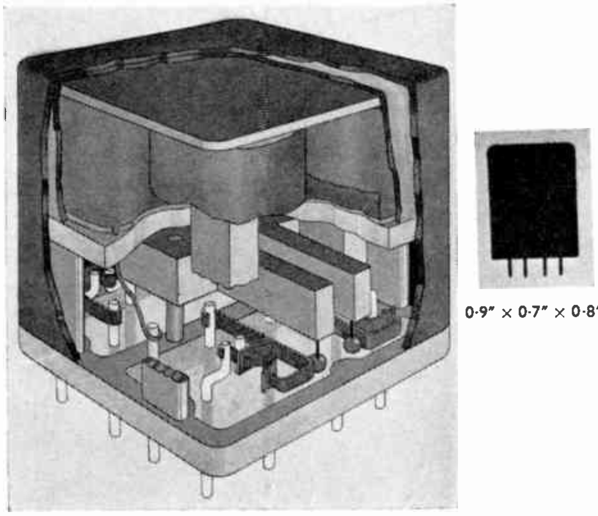


Fig. 4. Four-coil four-contact stabilized relay. Comprising two assemblies of the type used in the relay shown in Fig. 3 mounted together.

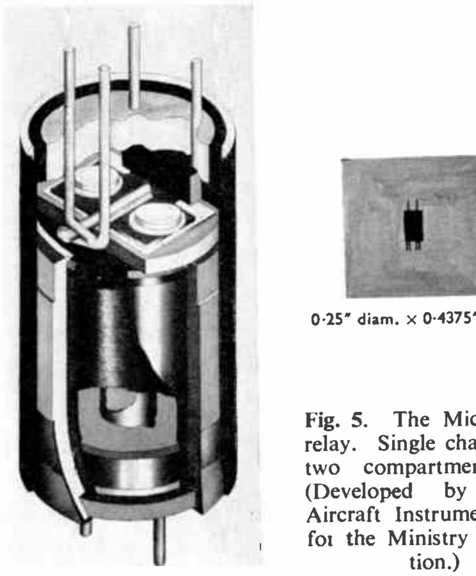


Fig. 5. The Micro PCR relay. Single changeover, two compartment type. (Developed by Smiths Aircraft Instruments Ltd. foı the Ministry of Aviation.)

high *g* vibration and acceleration in the earlier design mentioned. These constrained springs withstand 100 *g* vibration over the frequency band 30–2000 c/s with contact pressures of 10 grammes. In a polarized version of this relay orthodox split cantilever springs are used, and greater sensitivity is obtained due to the introduction of the permanent magnet.

A third form of this relay comprises two of either of the former relays mounted side-by-side on a single header. The armatures are then linked together with a "stirrup-and-bar" mechanism in such a way that the armatures not only rotate in opposite directions, but cannot rotate in the same direction. This assures freedom from interference from torsional vibration. This relay is shown in Fig. 4.

The four coils permit a parallel/series connection which allows for a combination of short circuit connection and disconnection of three coils yet still maintaining operation of the relay on the remaining coil. This increased coil reliability is advantageous.

4. Micro Relay

This relay is scheduled for use in the multiplex automatic landing and throttle control system of the DH 121 and therefore requires to be extremely reliable.

Where space is extremely important and contacts are better situated with minimum wiring length, several single changeover contacts of this, the smallest two-compartment relay in the world, can be used. It is comparable in size to switching transistors and is shown in Fig. 5.

It consists of a pot-type magnetic circuit with an armature "lid" suspended on a short flat return spring which permits pivoting of the armature and moving contact to perform the changeover function.

The magnet is a separate sealed assembly, with the pole-face sealed by a soldered-in brass disc, and the lead-outs fixed with epoxy resin before attachment to the glass/metal header.

Gold-plated platinum-rhodium wire contacts, sealed into another glass/metal header, are formed to the correct gap. With the armature in position on the header pins, the closed contact is adjusted to give 10 grammes pressure.

High rigidity to combat vibration is obtained in any but the operating plane, along the axis of the relay, which is provided for by the high force/mass ratio of the armature of about 300:1. This results in a high *g* relay by virtue of the small weight of the moving parts, which is 0·047 grammes.

5. Conclusions

The reliability factor for these relays, or any other, has not yet been established. An automatic test console is nearing completion which will permit rapid testing with complete monitoring of contact behaviour throughout the tests. The results from this machine will be described in a forthcoming paper.

6. References

1. Specification DEF-5011 "Climatic and Durability Testing of Service Electronic Components".
2. Specification RCS.165 (DEF-5165) "Relay Armature Sealed".
3. Joint Service Standards for Radio Components.

Manuscript first received by the Institution on 30th November 1960 and in revised form on 28th February 1961. (Contribution No. 30.)

THE 1961 ROYAL SOCIETY CONVERSAZIONE

This year's Conversazione of the Royal Society, held in its Rooms at Burlington House, London, on 11th May 1961, included the customary exhibition of scientific research work. A representative of the Institution was invited to attend a preview of the exhibits and noted a number of items of interest to the radio and electronic engineer.

Two of the exhibits will be particularly familiar to members of the Institution since they have been the subject of meetings held in London recently. On 26th April, during a symposium on electronic counting techniques Dr. E. Franklin of the Atomic Energy Research Establishment, Harwell, presented a paper on "A Self-indicating Magnetic Scaling System using Electro-deposited Nickel-iron Film Cores". In collaboration with Dr. J. Stephen, also of A.E.R.E., the techniques of manufacture and the applications of electro-deposited nickel-iron films were demonstrated at the Conversazione.

These films of a nickel-iron alloy (82% nickel, 18% iron), are between 8 and 60 micro-inches thick and are deposited on plane or cylindrical supports. During deposition a magnetic field is applied, which orients the magnetic film as it grows. The film has a rectangular hysteresis loop with two stable states for the remanent flux, providing two semi-permanent memory conditions; suitable pulses make the film switch from one memory state to the other, with a speed of switching between one thousandth and one tenth of a microsecond. The nickel and iron are electroplated simultaneously on copper and gold supports to form an alloy with the required characteristics.

A cylindrical film is used in the self-indicating magnetic memory scaler which is shortly to go into pilot production. Individual units of the scaler are "scales of ten", each unit using ten films and twelve transistors. The films are arranged in a circle and are magnetized with nine films in one direction and the tenth in the reverse direction. A compass needle in the centre of the circle points to the "odd" film, corresponding to an indication of zero. As counting proceeds, the "odd" condition is stepped on one position for each pulse and after ten pulses the "odd" condition is at the start again, and the pulse is passed to the next "scale of ten". When counting stops the compass needle of each scale of ten points to the "odd" film, thus indicating the number of pulses counted. The scaler is portable, battery operated with a very low power consumption and it can count faster than 1 Mc/s.

The so-called "radio pill" has several times been discussed at meetings of the Institution's Medical and Biological Electronics Group, notably as an example of the application of transistors. Modern methods of miniaturization have made it possible to construct radio transmitters fitted with measuring devices, the whole assembly being so small that it can easily be swallowed. It is thus possible to measure physiological variables such as pressure, temperature and pH, throughout the whole length of the gastro-intestinal tract without discomfort to the patient. A small loop aerial connected to a radio receiver picks up the signal from the radio pill and allows transmitted data to be recorded. Pressure-sensitive pills have been used in clinical research for some time, and typical recordings were demonstrated during the Conversazione.

The experience gained in the design and construction of radio pills has been used to construct other sub-miniature transmitters of rather greater dimensions but of much greater range. These transmitters are of fountain-pen size and are intended to be attached to the outside of the body. The range is several hundred yards and the transmitter can be modulated by a variety of transducers. The simplest version will accept voltages as low as 1 mV directly so that these can be used for the direct transmission of the cardiogram. Their development is part of a programme of "micro-telemetry" to enable measurements to be made on subjects during normal activity.

These exhibits were presented by Mr. A. M. Connell and Dr. E. N. Rowlands of the Gastro-Enterology Department, Central Middlesex Hospital, and Mr. J. McCall and Mr. H. S. Wolff of the National Institute for Medical Research.

The third exhibit was presented by Dr. M. B. Clowes and Mr. J. R. Parks (Graduate) of the National Physical Laboratory and dealt with the automatic identification of printed numerals. The central problem in this field is that of dealing with wide variations in the physical description of what is—to the human observer—the same pattern. There are variations of size, position, orientation, shape and intensity. The technique demonstrated overcomes the problems of variation in position and is potentially capable of overcoming the other difficulties.

Auto-correlation is often used to detect regularity or periodicity in a physical process. Thus, the average spacing of lines in a grating can be measured by superimposing two copies of the grating (or one copy and an image of it) and recording the variation of transmitted light as one copy moves relative to the other. This technique is the basis of the method of character recognition being developed at N.P.L. In order to obtain the necessary resolution more than two copies are required and in the two mechanical models displayed, there were three copies of each pattern.

The original character is in the form of a negative transparency, and the copies are formed by multiple reflections in a double-mirror system. For convenience in the optical design, straight lines are detected by a three-term auto-correlation function and curves by a five-term auto-correlation function. Two photomultipliers measure respectively the light resulting from three and five transmissions through the original. The orientation and position of the semi-reflecting mirrors determine the vector displacements, etc. The value of the auto-correlation functions is then measured to determine the identity of the character in the optical system.

Finally, though not electronic in character, an historical exhibit by the Science Museum consisted of a demonstration model of the original Jacquard loom control mechanism. This was invented in 1801 though it was first suggested as long ago as 1725: it is of course the basis of modern punched card selection and sorting techniques.

U.D.C. 631.395.625.3

Magnetic Recording Head Adjustment and Alignment Devices

By

R. B. DYER (*Associate*)†

Summary: This note considers the problems involved in mounting magnetic recording heads in a manner permitting ready adjustment to tape path, thereby avoiding losses inherent in non-adjustable mountings. Two constructions of the device are described, for professional and domestic grade heads respectively.

1. Introduction

The design and consistent manufacture of efficient magnetic recording heads, though the major problem, does not complete the case-history of the head assembly of successful tape recorders.

All modern recorders operate down to magnetic wavelengths of less than 0·0005 in; from the established formulae of separation losses it is evident that the maximum permissible separation error is of the order of a few microinches, and in practice contact is limited by the finite flexibility of the tape backing and the imperfect flatness of its coating.

In order to ensure that these are, indeed, the only influences causing random separations between tape and head, extreme care must be taken both with the tape transport system, and the exact alignment of head-faces; the former is not within the scope of this note, the latter requires amplification.

2. Head Mounting Considerations

Firstly, the surface on which the heads are mounted, either directly on the deck or on a removable head plate, cannot be rendered dead flat to within sufficiently fine limits, nor—a much more difficult problem—be guaranteed free from warping in service and under vibration and shock conditions, without involving disproportionate weight, bulk and expense.

Secondly, the mounting brackets themselves are subject to the same limitations, due to the difficulty of achieving exact perpendicularity to the head plate and determining the precise height of mounting-holes etc., above it.

Thirdly, the head itself, if mass-produced at reasonable cost, rarely has its working face truly parallel (or perpendicular) to its mounting face, its gap dead centrally positioned, nor the heights above deck and widths of spaces between its tracks accurate to within the necessary 0·001 or 0·002 in.

† Electric & Musical Industries Ltd., Hayes, Middlesex.

Any one or a combination of these minor dimensional faults can give rise to random separation losses, which, according to their nature and continuity, will appear when minimal as drop-outs, or when maximal as continuously impaired h.f. response.

The cost of accurate dimensioning to such fine limits, and its guaranteed maintenance during the life of a machine, is altogether disproportionate, and may be dismissed for any but the most expensive professional grade equipment. Means must, therefore, be sought to control head posture so as to compensate for these defects.

The methods adopted must be easily understood and handled by semi-skilled production and inspection staff; must remain self-locking as adjusted under normal transit vibration; and must be readily explainable in simple language and accessible for adjustment by service engineers, or in the last resort by the owner/operator.

Furthermore, it is increasingly found desirable, for the more ambitious designs, to incorporate easy interchangeability between standard $\frac{1}{4}$ in. and wider tapes. Since the lower edge of the tape has a fixed height because of the spool turrets, the mean height of the head must rise by half the difference in tape width, and provision must be made for this in the head mountings.

Yet in practice one rarely observes the provision of any adjustment other than for azimuth. Such height adjustments as are provided are usually of so crude a nature as to involve disturbing azimuth and other postural settings, are relatively inaccessible, and frequently invoke the unsound practice of relying on the tightness of something to prevent further movement under transit vibrational hazards.

Again, azimuth is usually adjusted about some pivotal point well outside the area of tape contact, presumably on the comfortable assumption that the magnitude of adjustment is too small to affect track alignment to any noticeable extent. This may well have been true with full-track heads ten years ago, is dubious for $\frac{1}{2}$-track heads, and for $\frac{1}{4}$-track heads with their narrow tracks, relative intolerance of tracking errors, and much greater angular adjustment for azimuth, this assumption is dangerous.

The ideal location for the azimuth pivot is the centre-back opposite the exact centre of the area of tape contact; thus rotation cannot disturb the mean track alignment, the worst that can happen is that the effective track-width will be increased by a factor proportional to the cosine of the angle which the gap—by manufacturing error—makes with the centre-line of the tape. This of course, means that the rear face of the head will take charge and that the working face must be dead parallel thereto. Such conditions can be guaranteed to within sufficiently fine limits only in a professional grade head, for which it then remains only to ensure exact adjustment of track height.

3. A Mounting for a Professional Grade Head

The track alignment assembly illustrated in Fig. 1 provides for this adjustment. The head is mounted by means of the azimuth pivot A which rotates freely in a transverse hole bored in the riser block E, being spring-loaded to eliminate sloppiness. The rear end of the pivot is screwdriver-slotted to enable the front end, which is threaded and fitted with a wide, flat, circular boss, to be screwed into a tapped hole in the centre of the rear face of the head.

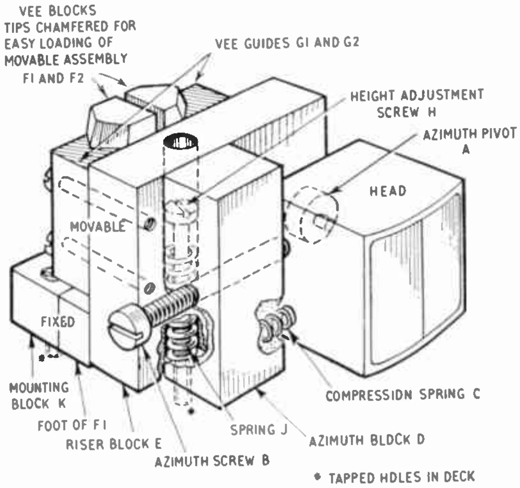


Fig. 1. The adjustable mounting for professional grade magnetic recording heads.

The azimuth block D is screwed to the riser block E, and azimuth adjustment is effected by screw B pressing on the upper side face of the head, opposed by spring C pressing on the lower side.

To the rear face of riser block E are secured two vee-guides G1 and G2, whose inner faces engage with the outer faces of the vee-blocks F1 and F2. F1 is secured to mounting block K which in turn is secured to the tape deck. F2 is inhibited from vertical motion, and permitted horizontal motion in one plane only, by means of horizontal hard steel pins driven into the one and sliding in the other very tightly, permanent lubrication being by molybdenum disulphide grease. They are also forced apart by compression springs recessed into each, so that engagement with vee-guides is very firm and free from all clearance-movement.

Consequently the whole main assembly, carrying the head, rides up and down relative to the deck in a manner entirely restricted to vertical motion with great accuracy, thus avoiding changes of azimuth setting and head posture relative to the tape.

The height adjustment screw, its head recessed in a counterbore in the top face of the riser block, runs right through into a tapped hole in the deck, opposed by spring J in a co-axial counterbore in the under face, thereby controlling the height of the assembly above the deck and thus relative to the tape.

Since all holes in the riser block are duplicated symmetrically about a transverse centre line, it may be reversed to the opposite hand without modification: thus the head may project at either the left or right hand end of the assembly as convenient.

4. A Mounting for a Domestic Grade Head

Reverting to domestic grade heads, we find that none of the necessary parameters may be controlled sufficiently at low cost, and the head adjustment assembly, illustrated in Fig. 2, was devised for ready adjustment and maintenance of control over all possible errors.

As it will be observed, the body of the device rides up, down, and around the central pillar. Height may be thus adjusted by pressure of nut A against spring B, and horizontal scan (to centralize the gap) by pressure from screw C working in a bracket integral with the baseplate, against bracket E, integral with the movable assembly, opposed by spring D.

Any out-of-parallel between the rear and working faces of the head may be compensated by pressure from screw F upon a lever rotating a transverse shaft through the movable block, thus varying the angular posture of the head. In this connection it should be noted that in theory this shaft should be exactly aligned both with the central horizontal plane of the head, and with the vertical plane of its front face, but this is impossible since the tape must necessarily occupy that location. Any alternative location to the

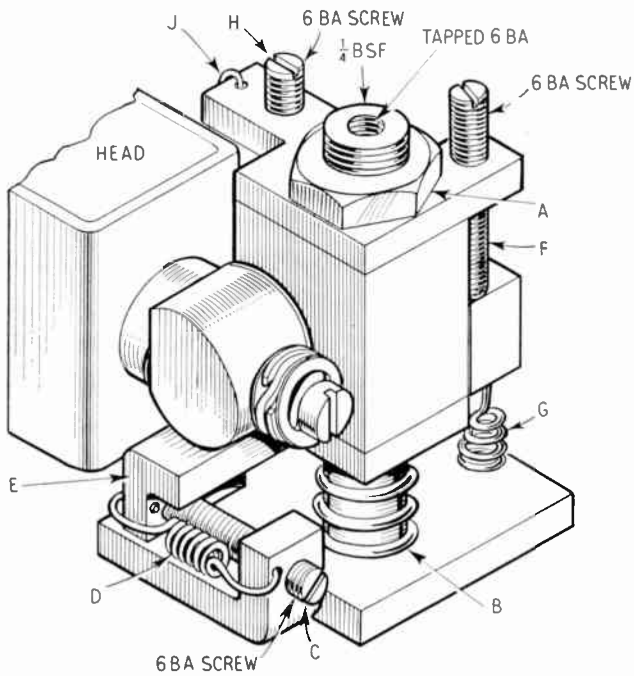


Fig. 2. The adjustable mounting for domestic grade heads.

A, B	Height adjusting nut and spring.
C, D, E	Horizontal scan adjusting screw, bracket and spring.
F, G	Out-of-parallel adjusting screw and spring.
H, J	Azimuth adjusting screw and spring.

rear of that vertical plane must of necessity cause postural changes to be accompanied by an undesired change of height, but as the latter is of small magnitude and well within the control range of the height adjustment nut A, the advantages of its design and location well to the rear greatly outweigh the obligation to effect a second minor compensatory readjustment.

Finally, azimuth may be adjusted by pressure from screw H opposed by tension-spring J, both acting upon a foot—not visible in this sketch—fixed to the bottom of the head and projecting to the left when viewed from the front.

5. Conclusions

Neither of the two devices described requires a generally high degree of manufacturing tolerance and finish. Both have proved in practice to be far less costly solutions to the head position problem than close control of the relevant parts.

Manuscript received by the Institution on 13th March 1961. (Contribution No. 31.)

Radio Engineering Overseas . . .

The following abstracts are taken from Commonwealth, European and Asian journals received by the Institution's Library. Abstracts of papers published in American journals are not included because they are available in many other publications. Members who wish to consult any of the papers quoted should apply to the Librarian, giving full bibliographical details, i.e. title, author, journal and date, of the paper required. All papers are in the language of the country of origin of the journal unless otherwise stated. Translations cannot be supplied. Information on translating services will be found in the Institution publication "Library Services and Technical Information".

THE ALCATRON

A French engineer has recently described a field-effect device of annular construction where the two grids have been separated. One of these grids—a small one—is used as a control grid. The second and larger one makes it possible to get rid of the dissipation power and modifies the equivalent circuit at high frequency. The first experiments made with such a device made from germanium have made it possible to go beyond the limits foreseen by earlier theories for the field-effect. These theories are rapidly examined and there is an explanation of why they do not apply directly to the present construction. The possibilities of further development for such a device are briefly sketched in relation to the existence of new semiconductor materials.

"Examination of a new tetrode field-effect device: the alcatron", J. Grosvalet. *L'Onde Electrique*, **41**, pp. 123–131, February 1961.

DECODING CIRCUITS FOR N.T.S.C. COLOUR

A recent German paper outlines the circuit and dimensioning problems appearing in the design of decoding facilities for colour television by the N.T.S.C. system, and proposes ways for their solution by reference to the example of a colour measuring demodulation. Particular attention is paid to the different methods of synchronous demodulation and their application to equi-band and I/Q receivers. Another new problem in the field of television engineering is automatic phase control of the colour subcarrier oscillator for which also various possible solutions are devised. A number of additional problems such as clamping and blanking of colour television signals, optimum matrixing of the colour difference signals, and measuring facilities as required for the operation of a studio type colour demodulator completes the survey of the present circuit techniques for N.T.S.C. colour demodulators.

"Circuit techniques of demodulators for colour signals in the N.T.S.C. system", F. Jaeschke. *Archiv der Elektrischen Ubertragung*, **15**, pp. 187–99, April 1961.

RADIO RELAY SYSTEM PLANNING

The ever-increasing requirements for long distance telephone and television circuits in the last few years, and the development of high grade f.m. radio systems designed to meet these needs, have contributed towards a large amount of the research done on radio propagation and radio system performance in recent years. Much information is presently available in the technical literature for engineering f.m. radio systems operating in the u.h.f. and s.h.f. bands and assessing the performance and reliability of such systems. An engineer with the Quebec Telephone Department has collected together information from a number of sources discussing different aspects of the system engineering, and giving the basic formulae and associated graphs. The calculated results from the formulae quoted are compared with measured values obtained with a radio equipment operating at a frequency of 6000 Mc/s.

"Planning of radio relay systems", K. Zement. *The Engineering Journal (Engineering Institution of Canada)*, **44**, pp. 45–55, April 1961.

VARIABLE CAPACITANCE AMPLIFIER DESIGN

There are two possibilities for tuning a variable-capacitance straight-through amplifier over a wide range of frequencies: tuning of the signal frequency and the pump frequency in the same sense with a fixed idle frequency, or tuning of the signal frequency and the idle frequency in opposite senses with a fixed pump frequency. It is shown in a recent German paper that tuning of the signal frequency and the idle frequency in opposite senses offers more advantage.

A laboratory model of a tunable variable-capacitance straight-through amplifier is described which covers a tuning range of 480 to 750 Mc/s. By tuning capacitors ganged in opposite senses the resonance frequencies of the signal circuit and the idle circuit are varied in such a way that their sum always equals a constant pump frequency of 1900 Mc/s. The bandwidth and the noise figure of the amplifier are investigated. Results of measurements on the laboratory setup are presented and shown to be in good agreement with theory.

With a bandwidth of 10 Mc/s an improvement of the noise figure obtained by a factor of 4·5 to 2·2 over the noise figure of a conventional triode amplifier in a grounded-grid configuration.

"A tunable variable capacitance straight through amplifier for the u.h.f. range", K. Bomhardt. *Archiv der Elektrischen Ubertragung*, **15**, 153–60, March 1961.

TRAVELLING-WAVE PARAMETRIC AMPLIFIER

The effect of various circuit parameters on the bandwidth and insertion loss of a stripline filter used as a slow-wave structure for travelling wave parametric amplifier are investigated in a recent Canadian paper. An amplifier has been designed and a description of an experimental model and its performance are given. Some of the disadvantages of using variable capacitance diodes in standard crystal packages are pointed out.

"Design and performance of an experimental L band travelling wave parametric amplifier", J. K. Pulfer and B. A. Howarth. *Transactions of The Engineering Institute of Canada*, **4**, No. 3, 1960.

SUBJECT INDEX

Papers and major articles are denoted by printing the page numbers in bold type

INDEX OF ABSTRACTS

This index classifies under subject headings the abstracts published through the volume in "Radio Engineering Overseas", synopses of Convention papers and reports of papers read before Local Sections.

JOURNALS FROM WHICH ABSTRACTS HAVE BEEN TAKEN DURING THE FIRST SIX MONTHS OF 1961

Acta Technica (Hungary)

Annales de Radioéléctricité (France)

Archiv der Elektrischen Ubertragung (Germany)

Canadian Electronics Engineering

Electronic Applications (Holland)

Electro-Technology (India)

Electronique et Automatisme (France)

Engineering Journal of Canada

Journal of the Institute of Electrical Communication Engineers of Japan

L'Onde Electrique (France)

Mitsubishi Denki Laboratory Reports (Japan)

Nachrichtentechnische Zeitschrift (Germany)

NEC Research and Development (Japan)

Periodica Polytechnica (Hungary)

Philips Technical Review (Holland)

Proceedings of the Institution of Radio Engineers Australia

Review of the Electrical Communication Laboratory, N.T.T. (Japan)

Slaboproudy Obzor (Czechoslovakia)

Transactions of the Engineering Institute of Canada

INDEX OF PERSONS

Names of authors of papers published in this volume are indicated by bold numerals for the page reference.
Authors of Brit.I.R.E. papers which are given in abstract form are noted by A.
Contributors to discussion are indicated by D.
Biographical references are indicated by italic numerals.

(NOTE: Authors of papers abstracted from other journals are not included here.)