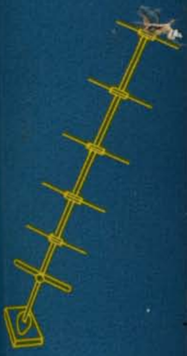
RADIO
ASTRONOMY
FOR
AMATEURS

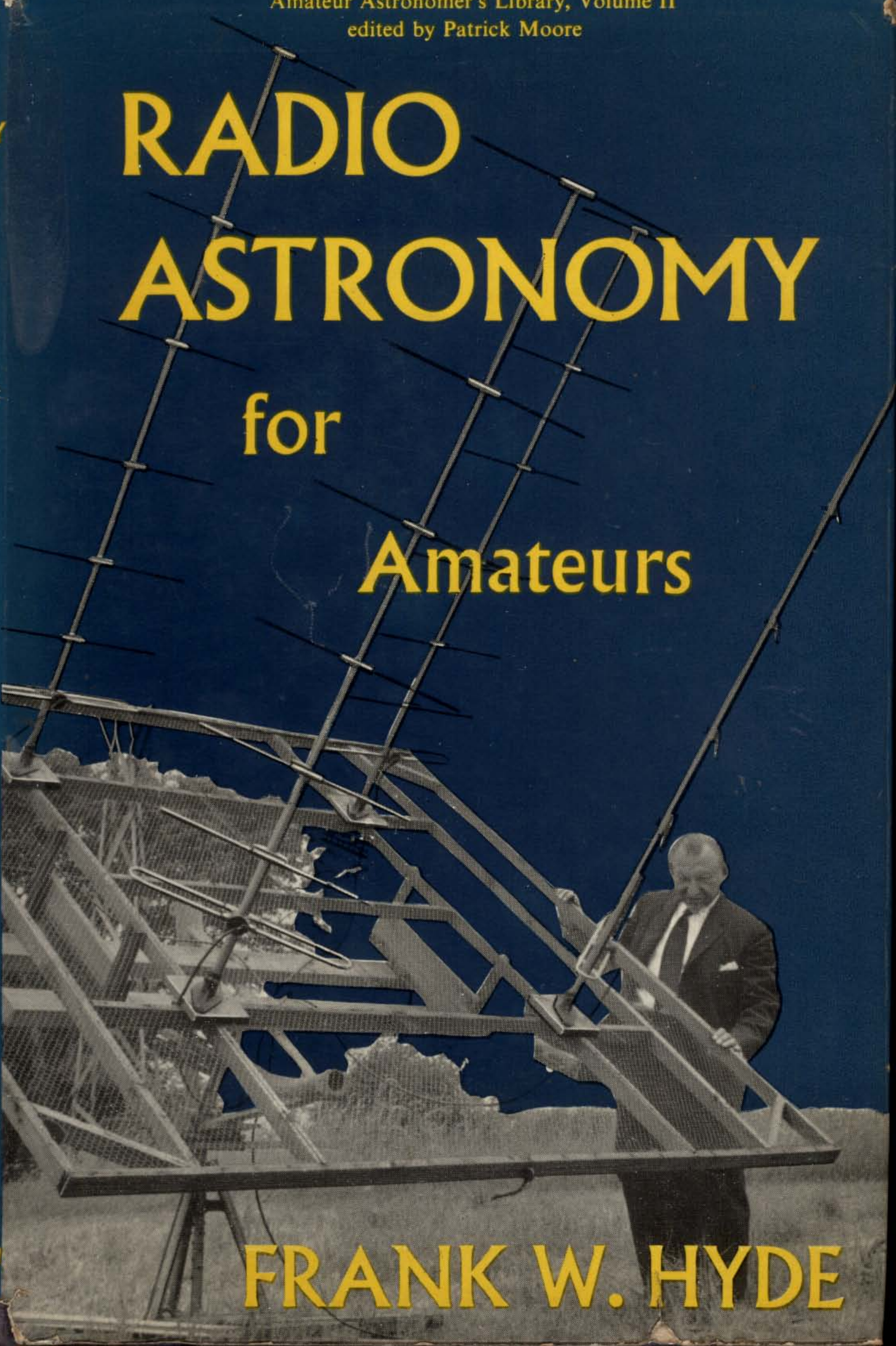


FRANK
W. HYDE

LUTTERWORTH

Amateur Astronomer's Library, Volume II
edited by Patrick Moore

RADIO ASTRONOMY for Amateurs



FRANK W. HYDE

Amateur Astronomer's Library

Volume Two

edited by Patrick Moore

It is only in recent years that radio astronomy has come into prominence, but although we are familiar with the great radio telescopes of the world, it may come as a surprise to many to learn that the construction and operation of equipment that can do very useful work is within the scope of the amateur at a modest outlay.

The purpose of this book is to bring together information on astronomy and radio astronomy for those who wish to venture into this exciting field of amateur activity.

Besides describing radio astronomy and its history, chapters are provided for those who have a working knowledge of electronics but are now coming to astronomy for the first time. Similarly, for those whose interest has been entirely in the field of optical astronomy, chapters deal with electronic fundamentals. The text is illustrated with 13 photographs and over 70 line drawings, including circuit diagrams.

Mr. Hyde, who has built five radio telescopes for his own use, is an acknowledged expert on the subject, a Council Member of the British Astronomical Association and Deputy Director of the Radio and Electronic Section. His experience in the field of amateur radio telescope construction gives him unrivalled knowledge of the practical problems amateurs may meet, so this book, the first to present this new approach to astronomy as a hobby, will be widely welcomed.

25s. net

RADIO ASTRONOMY FOR AMATEURS

by

FRANK W. HYDE, F.R.S.A., F.R.A.S., M.S.E.

Line diagrams by
RALPH GORING

LUTTERWORTH PRESS
LONDON

First published 1962

PRINTED IN GREAT BRITAIN BY
COX AND WYMAN LTD., LONDON, READING
AND FAKENHAM

CONTENTS

LIST OF CHARTS

LIST OF TABLES

LIST OF PLATES

LIST OF FIGURES

FOREWORD

by

PATRICK MOORE

Editor of the Amateur Astronomer's Library

ASTRONOMY is an expanding science. It has made remarkable progress during the past half-century, and our knowledge of the universe has been improved out of all recognition. This is due only partly to the construction of great new telescopes, such as the Palomar reflector, combined with new photographic techniques. Space-probes have arrived to extend our field of research; but from a purely astronomical point of view, the development of radio studies is probably much more important. Nowadays we hear so much about radio astronomy that it is not always easy to remember that it was completely unknown less than thirty-five years ago.

Most people have seen photographs of the 250-foot "dish" at Jodrell Bank; and who has not heard of that great radio-astronomy pioneer, Professor Sir Bernard Lovell? What is not so generally known is that amateurs, too, can make themselves useful in the overall programme. Moreover, they can do so without the expenditure of a large sum of money.

This is the theme of Frank Hyde's book. He is well equipped to write it, since he was himself a pioneer, starting with a small installation in the back garden, and now has set up a radio astronomy observatory which is not only the most elaborate in the hands of an amateur in Britain, but probably in the world. The reader will find here a description of the work which can be done, and will be given instructions as to how to undertake it. The full picture has not previously been given in collected form, and it will almost certainly stimulate many amateurs to enter the field – with results which will be beneficial not only to themselves, but also to astronomical science.

I cannot claim to have anything but the most rudimentary knowledge of electronics, and it may therefore be asked why I am writing the Foreword to a book in which electronics plays so vital a part. The answer is that as an amateur optical astronomer, I realize the importance of combining all methods of research in the amateur as well as the professional field; only in this way can we make full use of our opportunities. There are few serious amateur radio astronomers as yet, but in the coming years there will be many more, and what has been written in the following pages of this book will serve as an invaluable guide.

July 31, 1962 PATRICK MOORE

INTRODUCTION

THE PURPOSE OF this book is to bring together information on astronomy and radio astronomy for those who wish to venture into this exciting field of amateur activity.

Chapters are provided for those coming to astronomy for the first time, such as people whose previous interest lies in electronics. Similarly, for those whose interest has been entirely in the field of optical astronomy, chapters deal with electronic fundamentals.

As there is a wealth of information in the many books that have been written on both the subject of astronomy and electronics, no attempt has been made to cover in detail either of these subjects and a selection of the publications available to readers is given in the bibliography.

Those who are already well versed in optical astronomy need therefore only turn their attention to chapters which contain information new to them. To readers already well acquainted with the electronic sphere the elementary chapters on electronics will serve to show the extent of the ground to be covered in the sections dealing with the construction of equipment.

I hope that this book will serve as a guide in bringing together the aspects of both radio techniques and astronomical observations. If it stimulates individuals not previously active in this field, or brings together enthusiasts of astronomy and radio in joint endeavours, the efforts that have gone into the preparation will have achieved their object.

FRANK W. HYDE

Beacon Hill,
St. Osyth, Essex.
April 1962.

Chapter One

WHAT IS RADIO ASTRONOMY?

Radio astronomy presents a new extension of the frontiers of astronomy and once again enables man to step forward in his endeavours to solve the mysteries of his environment. The techniques used are those of radio communication and the rapid advance of work in this field of electronics has provided a powerful tool for the observation of the universe. It is not, therefore, something new which replaces optical astronomy but is rather an adjunct to it.

The early astronomers had only their eyes as instruments of observation. Such, however, is man's natural curiosity that before long he had added artificial aids to those supplied by nature; he devised instruments, which though they did not enable him to see more than he could with the naked eye, made it possible to combine visual observations with measurement of position and direction. Thus man was able to obtain a limited understanding of the sky above him. In various parts of the world there are monuments in stone to man's attempts to record his measurements. Typical among such observatories is that of Jhai Singh in India; this observatory contains instruments of stone which to this day can be used for astronomical observations of considerable accuracy.

The advent of the optical telescope so widened man's view of the universe that many curious notions had to be discarded. As time went by, optical instruments became more and more complex until the limitation of the atmosphere made further progress difficult. It is perhaps a little ironic that the very medium which makes life possible for man on earth should be the means of restricting his view of the universe. However, this optical window through which extra-terrestrial objects could be observed had served astronomy well, information built up over the ages has enabled astronomers to make calculations and speculations on the past history of the universe; and, at the same time, make reliable predictions of the future movement of many celestial objects.

Optical telescopes increased in size and complexity and much sophisticated ancillary gear was added to a telescope to enable more reliable information to be obtained. Photographic techniques advanced so far that optical astronomy reached a stage where the elimination of errors became more and more important.

Just at the time when the limits of optical observation became more and more frustrating a new window was opened on the universe – this was the radio window. It was discovered by Jansky in America that at certain radio frequencies he was able to receive radiations which must have come from extraterrestrial sources. The radio window through which Jansky looked upon the universe is some way removed from the optical window in the electro-magnetic spectrum. The electro-magnetic spectrum comprises all the radiations known to science. The optical part of this spectrum comprises the region between the colours red and violet which are detectable by the naked eye. This portion of the spectrum comprising the visible light has become known as the visible spectrum. To observe any other part of the electro-magnetic spectrum special techniques are used depending upon which part of the spectrum is involved.

The radio spectrum extends from the frequency of 30,000 megacycles, or 1 centimetre wavelength, to a few cycles and many thousands of metres wavelength. The radio window through which we look at the universe, however, is not so extensive as this. It may be said to begin at 30,000 megacycles and extends as far as 10 megacycles or 30 metres wavelength. The radiations of celestial bodies appear in all parts of the electro-magnetic spectrum, and while some radiate more in the optical section of the spectrum, such as the Sun, others may radiate more in the radio section.

There is a difference between radio waves and light waves, although they all belong to the same electro-magnetic spectrum. Radio waves may be generated by electric currents directly, whereas light waves are emitted only by molecular or atomic processes. Radio waves may be used to detect the existence of ionized gas, however rarefied, but this would be transparent to visible light. On the other hand, ordinary clouds will prevent

visual observation of celestial bodies, but radio waves find clouds quite transparent. In fact, one of the chief advantages of radio astronomy is that it is largely independent of the weather. This is extremely important because its enables observations to be carried on over very long periods without any break due to obscuration of scene. Thus we now have two windows through which we can observe the universe.

Here is perhaps the point at which we should compare the optical telescope with a radio telescope. This will give a clearer picture of how they are related one to the other.

An optical telescope of the reflecting type is shown in Fig 1*a*, and Fig. 1*b* shows the radio telescope. The optical telescope receives light waves which are reflected by the primary mirror to a secondary mirror or prism, and thence to the eye-piece where the image may be examined. In the radio telescope the radiations are received at the reflector, collected by the aerial, passed to the receiver and thence to a recording device which may be an oscilloscope, a tape recorder or a pen recorder. The analogy is thus a very close one; in the optical telescope the mirror receives the light waves, and in the radio telescope the reflector and the aerial receive the radio waves. The magnification of the optical telescope is determined by the ratio of the focal lengths of the eye-piece to the primary mirror; the magnification of the radio telescope is determined by the gain

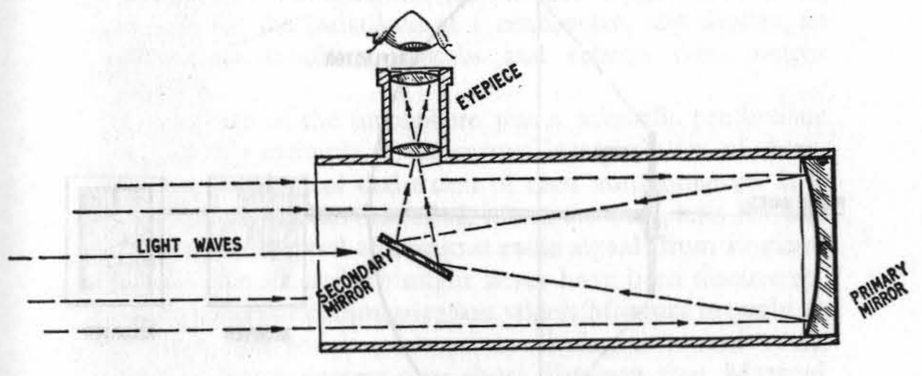


Fig. 1a Optical reflecting telescope

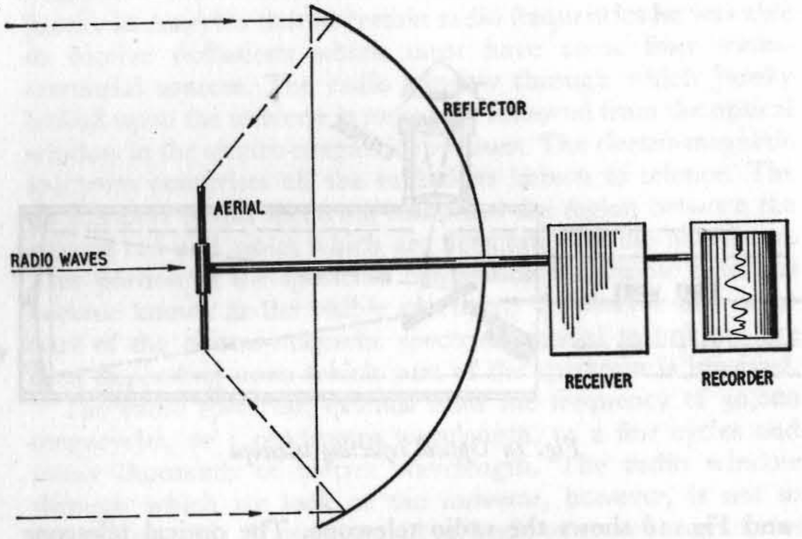


Fig. 1b Radio telescope

of the receiver and the gain of the aerial. The optical telescope enables the image to be studied by means of an eye-piece, whereas the radio telescope does this by direct audio recording, or by observing an oscilloscope or by making a permanent record with a pen on a paper roll. The recorder is therefore the eye-piece of the radio telescope. The pen recording has the greatest advantage since it is a permanent record. Here the analogy with photographic records made by the optical telescope will be apparent. These records can be studied in detail at leisure just in the same way that photographs made by optical telescopes can be studied. New information which becomes available can be checked against previous records; one particular example of this occurred when Burke and Franklin in America first detected radiations which they assumed to come from Jupiter. Shain in Australia then looked back on his records made several years previously to see whether he could find confirmation of these observations, and in this he was successful. The essentials of a radio telescope are then: (i) an

aerial, (ii) a receiver and (iii) a recording device. With comparatively simple equipment it is possible to examine extra-terrestrial radiations.

We have said that these radiations come to us through the second window – the radio window. Just as the optical window is limited in its extent, so is the radio window. At wavelengths shorter than 1 centimetre the window becomes obscured. The reason for this is that the radiation is absorbed by oxygen and water vapour. As the wavelength gets longer we reach a point at about 30 metres, where the ionosphere begins to take over. Except at special sites, or the advent of certain special conditions of the ionosphere, very little work can be done in radio astronomy at longer wavelengths. Just as the oxygen and water vapour absorbs the radiations at 1 centimetre, and shorter, so the ionosphere modifies, absorbs and reflects these longer wavelengths.

The existence of the ionosphere was a scientific prediction: it is yet another example of the correct interpretation of observations and the logical deduction of their consequences. It is also a warning that our knowledge is never complete, for had not Marconi attempted to transmit radio signals from England to America, the ionosphere might never have been discovered. The new method of communication which Marconi brought to this country, and which we know as wireless communication, met with so much success over short distances that Marconi was encouraged to try much longer distances. He was advised by scientists that this would be impossible; they said that owing to the curvature of the Earth, when the waves reached a certain distance they would go straight out into space. Marconi, however, was a practical engineer and was determined to try the experiment. This is the secret of the experimenter, the man who will try things even though others tell him that it is impossible. Marconi, with his colleagues who shared his enthusiasm, was able to establish a transmitting station in England and a receiving station on the other side of the Atlantic. His faith in the project was justified, and when the first faint sounds of the letter "S" were received he had posed a new problem for scientists. Not unnaturally this caused a considerable stir. Two brilliant men, Kenelly in America and Heaviside in England, developed theories to show how this transmission might be possible. Oliver Heaviside in England was a man who has brought great brilliance in transmission theory to Post Office communications. He and Kenelly developed theories to show that a layer of charged particles must lie high up in the atmosphere. They saw this as a mirror-like shell around the Earth; radio waves they believed would go from the Earth up to this reflecting layer and back again to Earth. Many people became involved in this new method of communication, but the manner in which the transmission was effected remained theoretical until 1924. At this time the long-wave programme of the BBC, then known as 5XX, was found to have a background which was identified as the programme from Radio Luxembourg. At certain times these two signals were received together and it was immediately inferred that this was due to the reflecting layer. Professor Appleton set up an experiment to measure the delay between the direct waves from the radio transmitter and the wave covering the long distances up to the ionosphere and down again, so he, working with BBC engineers, established the existence of this particular layer. The difference in time between the two paths from the transmitter to the receiver enabled them to measure the height of this reflecting region. This we know today as the "E" region.

There are several different layers of the ionosphere; some which remain more or less permanent, but others which vary from time to time. The lowest of these, known as the "D" region, is at a height which varies between 35 and 50 miles. This is the region where the atmosphere is relatively dense and the atoms are broken up into ions by sunlight, but because of the density of the atmosphere they quickly recombine. The amount of ionization, therefore, depends directly upon the amount of sunlight, so we find that "D" region density of ionization is greatest at noon and tends to disappear at sundown. When radio waves enter the "D" region the electrons are set in motion and because of the high air density, collisions between the particles are very frequent; a good deal of the energy is therefore given up as heat. The number of collisions depends upon the distance that an electron can travel under the influence of the wave; this distance is dependent upon the frequency, because during a long wave the electron has time to move further before the direction of the field reverses and sends it

back again. Should the frequency be low enough, the collisions between the particles will be so frequent that all the energy of the wave will be absorbed in the "D" region. The "D" region, therefore, is our limiting side of the radio window, for at frequencies below 10 megacycles it tends to become opaque. This is confirmed by what we know of radiations from the Sun, for we now believe that it is the "D" layer which is largely responsible for radio fade-outs during solar flares. Under these conditions the "D" layer has become so dense that communication between one point and another at selected frequencies is not possible.

The "E" region of the ionosphere, which is at a height of about 60 miles, is fairly well understood; it is created by the ionization of oxygen molecules by the effect of sunlight: it is at its maximum around noon and falls off rapidly after sundown. The ions and electrons recombine so that at midnight the ionization is at a minimum, but it increases again rapidly at sunrise. This layer absorbs energy from low-frequency waves at the time of maximum ionization. During the night the reduction of ionization is irregular, and this results in scattered layers which are somewhat like ionized clouds. These areas are known as "sporadic E".

The highest levels of the ionosphere are known as the "F" region. This is the result of bombardment by ultra-violet and x-rays which ionize the oxygen atoms and nitrogen molecules. It is most dense and nearest the Earth at midday in summer when the Sun is high in the sky and the Sun's radiation can penetrate further into the Earth's atmosphere. This "F" region extends to a height of some 300 to 400 miles. Ultra-violet light allows the outer electrons to be detached from oxygen atoms and become free; thus the amount of atomic oxygen is increased and becomes more abundant than molecular oxygen. These facts are summarized in the diagram – Fig. 2.

The ionosphere has been extensively studied by radio methods, one of which is the measurement of echoes from meteor trails that revealed there were very strong winds at high altitudes. The presence of these winds had been suspected by studies of the movement of noctilucent clouds. These are clouds in the region between 50 and 60 miles above the Earth, and are just visible on some clear nights. The winds are comparable

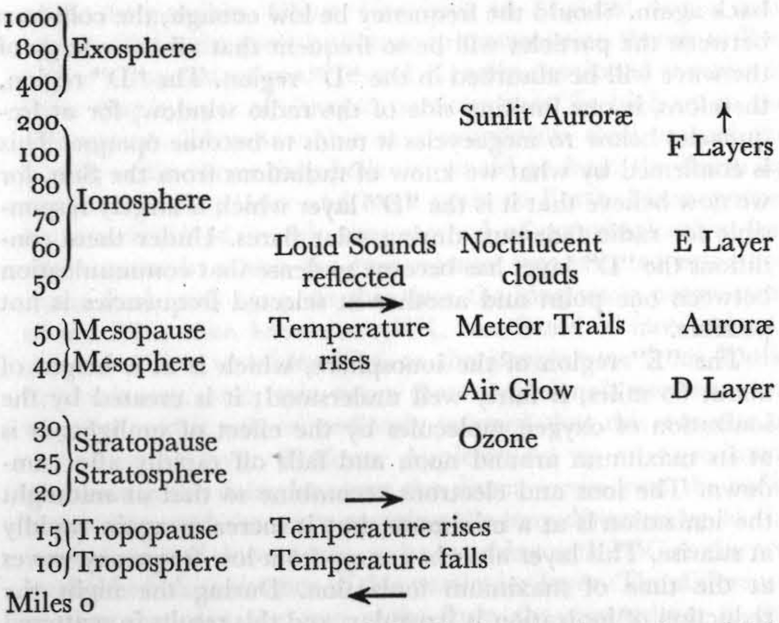


Fig. 2 Chart showing the extent of the Earth's atmosphere and associated phenomena

with those found near the surface of the Earth, and they may be of the order of 60 to 70 miles per hour, but they have the peculiarity of not only rapidly changing speed, but also they have unpredictable movements. At levels separated by only a few miles there may be winds blowing in opposite directions. There is, however, a definite sequence apart from these unpredictable movements: at twelve-hour intervals they blow clockwise round the compass, in summer and winter they tend to blow towards the east, but in the spring and the autumn they blow to the west. Our knowledge of this area is largely deduced from measurements made of meteor echoes.

We have already referred to the ionized clouds in the "E" region, but these clouds are most prominent in the "F" region. It is these clouds which are responsible for the scintillation of radio sources. Just as ordinary stars appear to twinkle or scintillate due to disturbances of the atmosphere, so do the radio sources scintillate as a result of the movement of these clouds. The passage of the clouds between the source and the radio telescope causes the level of the radiation received to vary. Extensive investigations of this phenomena were made in New Zealand and Australia by Stanley and Slee. These two radio astronomers made recordings at two different points simultaneously. The pattern of the fluctuations was quite different at each of the two sites, and from these observations it was quite obvious that the effects must be of terrestrial origin. If the variations had been due to changes of radiation from outer space both stations would have shown similar fluctuations.

Investigations were also made at Cambridge and Jodrell Bank, and again it was confirmed that these scintillations were of terrestrial effect. It could be shown that they occurred high in the Earth's atmosphere where the ionized clouds are found.

These clouds are extremely large; they vary in size from 2 to 6 miles and move at terrific speeds. In a few cases their size may be as great as 60 miles across; they are usually sausage-shaped and they are moved about by a kind of "wind". Now this is not wind as we ordinarily understand it where air is moved about by the change of pressure from one place to another; these "winds" are due to the effect of the Earth's magnetic field. The speeds of these clouds average some 500 miles an hour, but extreme cases have been noted where the speed is of the order of 2,000 miles an hour. There is a regular sequence about this "wind", for it blows towards the west in the evening and swings round about midnight to blow towards the east. The regulation of these clouds is controlled by the Earth's magnetic field, and when there are large changes in the Earth's magnetic activity caused by particles ejected from the Sun, great increases are shown in scintillation records. It is not clear exactly why these clouds form, but it has been suggested that some of it may be the result of the impact of material which comes from outer space colliding with the ionized particles. There are, however, a number of things which need explanation. One of these is the fact that scintillations vary from day to day, even though extensive disturbances are not present. Sometimes scintillations are not apparent at all. One suggestion has been made that they may be due to eddies in the "F" layer of the ionosphere, but it is difficult to understand how this is possible, as the gas at this level is too tenuous for it to have very much effect.

There is another characteristic of the ionosphere which is of importance to the radio astronomer: this is known as Faraday rotation.

Michael Faraday, during his experiments, showed that light could be rotated by means of a magnetic field. We find that the ionosphere will also cause radio waves to rotate, and this is brought about by the combined effect of the Earth's magnetic field and the density of ionization. Faraday effect enables us to determine the extent of ionization and the density of the electrons. This has been done in the case of echoes from the Moon; a radar pulse is sent to the Moon and during its journey to the Moon and back again it is twice rotated. The amount of rotation indicates the number of electrons present, which makes it possible to predict the height at which various frequencies will be absorbed. By this means we can forecast the highest frequency which can be effectively used for communications. Since we know that the density of the ionosphere varies during different parts of the day we are able to decide when is the best time to make our observations.

These, then, are the limiting effects of the radio window through which the radio astronomer looks upon the universe. It will be clear that only a portion of the electro-magnetic spectrum is visible through this radio window. Nevertheless, it has extended astronomy in general far beyond the horizons set by the optical window. While optical astronomy deals with that which is visible and can be photographed, radio astronomy deals with what lies between the stars as well as that which is visible. By the use of radio astronomy we can detect the existence of gas in the universe which cannot be seen in any other way. Radio astronomy can take us to farther horizons than optical astronomy. With the two techniques combined we have a much greater opportunity of discovering what is going on outside the Earth. This is an exciting and fascinating field for the amateur. Though in few cases will he be able to erect the extensive aerial systems, such as the professional observatories use, yet there is much that he can do. Just in the same way that a 6-in. reflector will enable an amateur to do useful work in certain spheres of optical astronomy, so will the amateur radio astronomer be able to make his contribution. The average garden is some 60 or 70 ft long and with this space available the amateur can work with a fairly high degree of resolution.

Resolution is one of the problems of radio astronomy. The resolution of a telescope is its ability to distinguish between two close objects. The Jodrell Bank radio telescope with its bowl 250 ft in diameter has a resolution at a wavelength of 1 metre of about 1 degree. If we compare this with the 200-in. reflecting telescope at Mount Palomar, we are better able to appreciate the difference in resolution. The 200-in. telescope at Mount Palomar has a resolving power of about one-tenth of a second of arc; in the case of the Jodrell Bank radio telescope the best that it can do in distinguishing between two sources requires that they have to be 36,000 times further apart than those that could be distinguished with the optical telescope at Mount Palomar.

There is another difference too between the radio telescope and the optical telescope, and this is in the way in which it "sees"; with the optical telescope we can produce an actual image of the whole of the part of the sky at which the telescope is pointed – we can record this entire area in one operation on a photographic plate. In the case of the radio telescope we are dependent upon the voltages which appear at the aerial, and these we have to translate into a change of level at the receiver. In order to form a picture of the radio sky, therefore, we must scan it in strips; having done this, we can place all these strips together to produce our final picture. The type of picture that this produces is rather different from that of the photograph. A typical map made by a radio telescope is shown in Fig. 3. We are able to listen by means of a loudspeaker to the noise generated by these voltages that arrive at the aerial from outer space. When we do this it sounds very much like a radio set when it is not tuned to any particular station. There are, however, methods of using radio telescopes so as to overcome a good many of these difficulties, for example, even a simple radio telescope, which we call a radiometer, reveals that the Sun, from the point of view of the radio astronomer, is many times larger than that of the visible Sun. The optical telescope is only able to see the inner corona at the time of a total eclipse. The radio astronomer can examine this at any time. When we look

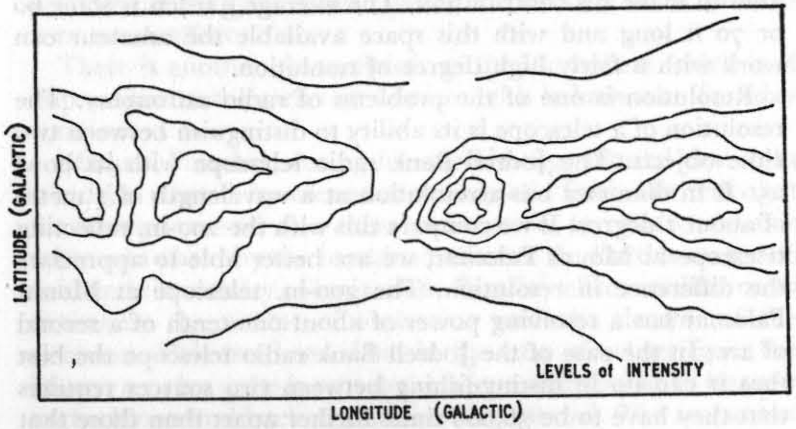


Fig. 3 Radio map of sky

up at the Milky Way we see many hundreds and thousands of stars, but we also see dark patches. The optical astronomer can see nothing in the dark patches, but the radio astronomer can often detect quite strong radiations from these areas. A combination, then, of a knowledge of astronomy and a knowledge of electronic techniques will provide the amateur with a new and exciting field of endeavour where his ingenuity can have full expansion.

Quite simple apparatus can be constructed which will enable the amateur to follow in the footsteps of the pioneering amateur, Grote Reber, who, searching for fresh fields to conquer, chose to look to outer space. The cost is no more than any other hobby and is essentially one where new techniques are continually needed. This is a field where those with a flair for building their own apparatus can shine. The surplus radio equipment that is available is well suited to this purpose and the construction of aerial systems is no more difficult than the construction of an optical telescope.

Chapter Two

THE HISTORY OF RADIO ASTRONOMY

IT IS NECESSARY for the proper understanding of a subject to know something of its history and background. Radio astronomy may well be said to have had its beginnings in the years between 1928 and 1932. In those days it was not called radio astronomy; its christening with this name did not come until some years later, after World War II.

In 1927 a student graduated from the University of Wisconsin and joined the staff of the Bell Telephone Laboratories. His name was Karl Guthe Jansky. The company posted him to their field laboratory at New Jersey and there assigned him the task of investigating the causes of interference on transoceanic telephone links. During the course of this investigation he noticed certain peculiarities about the noise in his receiver. His investigations covered a very wide range of frequencies but at one particular section of the spectrum, between 20 and 21 megacycles, there was a peculiarity which attracted his attention. Jansky observed that there was always a rise in noise when the aerial was pointed in a certain direction. He developed a large steerable aerial which he called his "merry-go-round". In those early days techniques in this field were entirely new and there were many problems to be solved. One of the main difficulties was to make a receiver which was sufficiently sensitive. After many trials and modifications to his equipment he was able, in 1929, to begin some positive observations.

At first Jansky thought the radiations that he was receiving were coming from the Sun. As time went by he found that the point in space from which these radiations appeared to come moved further and further away from the Sun until he was hearing them in the middle of the night. He finally came to the conclusion that these radiations must come from outside the Earth and from stars in the direction of the Milky Way. Jansky made two suggestions as to the origin of these radiations, both of which are of great interest today. He supposed that some heavenly body must be the cause of the radiations; such a body would be emitting radio waves in a greater amount than it was in light and heat. This was his first suggestion. The second suggestion was the result of his recognition that the sound of the radiation he received was very similar to that produced by thermal agitation in a resistor carrying current. He considered that a mechanism of this kind, where particles having electric charge were in constant thermal agitation such as in stars, would explain the origin of the radiation. The temperature required to produce this kind of radio emission would be in the region of 15,000 to 20,000 degrees Kelvin. There is, in fact, a considerable amount of interstellar matter distributed in the galaxy with an effective temperature of this order. Later observations have fully confirmed the correctness of Jansky's view.

One of the most dramatic confirmations of his reasoning about bodies of greater radio energy output in relation to the light energy output was the discovery of the source in the constellation of Cygnus. This has a radio emission 10^{-18} that of the Sun. It is now believed that a large part of the Milky Way is composed of sources of finite size like stars.

It is impossible to overestimate the importance of Jansky's work, for here was another of those remarkable instances which so often occur where work directed to one goal opened up a new avenue of exploration and pushed the scientific horizon outward still further. Jansky began work on a practical scientific project the object of which was the improvement of radio telephone communication. His ability to recognize something unusual brought to light a fundamental discovery in the field of pure science.

It is fitting to remember here that the last paragraph of Jansky's paper published in 1932 says that: "In conclusion, data have been presented which show the existence of electromagnetic waves in the Earth's atmosphere which apparently come from a direction that is fixed in space. The data obtained give for the co-ordinates of this direction a right ascension of 18 hours and a declination of -10 degrees." This was the part of the Milky Way towards the galactic centre. After the presentation of this paper he was transferred to other work and did not himself pursue the matter.

Although Jansky's discoveries made headlines and the sounds

which he recorded were even broadcast, the whole matter was soon forgotten and no serious attention was given to his discoveries by professional astronomers. It was left largely to the activity of amateurs to continue this absorbing study. One amateur in particular, Grote Reber, became interested in Jansky's early work.

Reber decided to work at very high frequencies because he assumed that then he would have a better opportunity of receiving signals. Though in theory his reasoning was sound it did not work out well in practice. The reason for this was partly due to the inefficiency of the apparatus at his disposal. However, he finally came to the conclusion that the frequencies that he had adopted were too high. He designed and constructed for himself a 30-ft diameter dish working at a frequency of 160 megacycles and with this apparatus he was able to repeat the work which Jansky had done. At this frequency of 160 megacycles the intensity of the radiations that he recorded were very much less than those noted by Jansky at 20 megacycles. Reber suggested the radiation was generated thermally by collision between free electrons and positive ions in the interstellar matter ionized by light. It may be truly said that Reber's 30-ft dish was the first instrument to be called a "Radio Telescope". With it he made a map of the radio sky; this map is still today a reasonably accurate picture of extra-terrestrial radiations. Reber had raised this fundamental discovery to the level of practical research. Once again the activities of the amateur scientist had made its mark upon the history of the human quest for knowledge and opened up an ever-expanding field of research. It is a great pity that the originator of this early work, Karl G. Jansky, did not live to see the extent to which his early discoveries had been pursued.

Other amateurs were also active in these fields and many radio enthusiasts had noticed the effect of the Sun on transmissions. In the years between 1932 and the beginning of World War II a number of publications and private communications have passed between amateurs, and some had appeared in scientific journals. Notable among these was a review of the activities in 1937 by D. W. Heightman. He indicated in this review, which was published in the *Wireless World*, that he had received signals which he supposed to be coming from the region of the Sun and suggested that they were due to charged particles ejected from the Sun and impinging upon his aerial. The first recognition of the reception of radio waves from the Sun was made in 1942 by Hey in England and Southworth in New Jersey in America.

It is worthwhile to remember that Sir Oliver Lodge in 1894, only 6 years after the discovery of radio waves by Hertz, had said: "I hope to try for an all-wave radiation from the Sun filtering out the ordinary well-known waves by a blackboard or other sufficiently opaque substance." He did, in fact, try some time between the years 1897 and 1900 because later in lecture he said, "I did not succeed in this endeavour, for a sensitive coherer in an outside shed, unprotected by the thick walls of the substantial building, cannot be kept quiet for long. I found the galvanometer spot of light liable to frequent weak and occasional violent excursions and I could not trace any of these to the influence of the Sun. There were evidently too many terrestrial sources of disturbance in a city like Liverpool to make the experiment feasible. I do not, however, think that it might not be possible for it to be successful in some isolated country place. But clearly any arrangement must be highly sensitive in order to succeed." It was more than forty years later that Southworth and Hey succeeded in detecting these radiations.

The association of radio fade-outs with the activity of the Sun was already under the notice of scientists. Studies were being made of short-wave radio fade-outs in connection with solar flares; terrestrial magnetic changes were being observed, as was also enhanced atmospheric static and sudden increase in noise-level.

The origin of this increase in noise level, however, had attracted the least attention. A paper did appear in Japan by Nakagami and Miya which described an experiment made in the 15 megacycle band to measure the angle of arriving noise. They found that this noise occurred only during the day, and that its direction of arrival was from a high elevation. They did not however note the significance of the elevation of the Sun and concluded that this noise probably originated in or near the "E" layer of the ionosphere.

In February 1942 a number of British army radar units operating at frequencies between 55 and 80 megacycles were

noticing severe interference. This was investigated by Hey and was described in reports during the war. These were, of course, restricted and not available to the general public. It was found that the direction of the interference was from the Sun and during this period a large sunspot was, in fact, visible on the disk. It was concluded that this sunspot, or something associated with it, was responsible for the activity. This was substantiated in Sydney by Pawsey, Payne-Scott and Macready in 1946. The relationship of solar noise and the sunspot activity was conclusively established.

In 1946 Ryle and Vonberg devised a new and much improved technique for the measurement of the size of radio sources. They developed an interferometer consisting of two aerials spaced a number of wavelengths apart in order to give a multi-lobe receiving pattern analogous to Michelson's method for measuring stellar diameters. They were, with this apparatus, able to measure sources which were less than 10 minutes of arc in diameter.

In 1947 Macready, Pawsey and Payne-Scott published the results that they had obtained in February 1946 giving similar results on the source sizes by using a cliff-interferometer which consists of only one aerial. This was situated 250 ft above the sea and utilized the reflected ray as well as the direct ray to produce a multi-lobe pattern. This system was similar to Lloyd's mirror interferometer.

In 1946 the polarization of these intense radio emissions from sunspot areas was first detected. They were found to be circularly polarized and these observations were confirmed by three independent groups in England and Australia. Since then both the thermal component from the quiet Sun and the non-thermal intense bursts from the active Sun have been the subject of detailed study. The extent of the corona of the Sun has also been investigated by radio astronomers. It is found that the radio corona extends far out beyond the visible corona, which can only be seen with a special telescope called a coronagraph, at the time of the solar eclipse. The corona sends out long streamers in all directions, and it is possible to measure the radio emission of these streamers but to a limited extent, since at the lower frequency the radiations cannot penetrate our ionosphere. As well as emitting radio waves the corona can defract or scatter radio waves. It follows from this that the corona will also occult or blot out a source of radiation.

At Cambridge in 1950 Ryle and Hewish attempted to do this using an intense radio source which lies near the ecliptic. This source was the Crab Nebula. It is a vast expanding mass of gas which is all that remains of a supernova observed by the Chinese some 900 years ago. It lies at a distance of some 4,000 light years from us. It was reasoned by the Cambridge workers that if the corona can defract and scatter radio waves, it should be possible to observe a reduction of the level of radiation received when the corona passed between us and the Crab Nebula. This in fact proved to be the case and independently in Russia the same experiment has been carried out with similar results.

These observations are now made regularly each year. This is one of the observations that can be made by an amateur and at the author's observatory it is a regular part of the programme of work.

The Crab Nebula has also been used to determine whether or not the Moon has an atmosphere. On January 24, 1956 this experiment was performed at Cambridge when the Moon passed in front of the Crab Nebula and the radio waves were observed to be occulted or blotted out for 59·6 minutes. Since the disk of the Moon took only 59·2 minutes to pass in front of the Nebula there was left an amount of 0·4 minutes to be accounted for. Interpretation of this result leads to the conclusion that what atmosphere there is, if any, on the Moon is about two parts in ten million million of the Earth's atmosphere. We do not yet know what the composition of that atmosphere could be.

Radio astronomy has been used to determine the distance of the Moon from the Earth by means of the echo technique. Attempts have been made to measure the distance of Venus. The study of thermal radiation from the Moon and the planets has also been undertaken. In 1945 Dicke and Beringer in America detected thermal radio emission from the Moon on a frequency of 24,000 megacycles. Piddington and Minnett in Sydney made the first extended series of observations of these emissions at various times in 1949. Observations of thermal emission at 3·5 centimetres has been undertaken in America

on Mars and Jupiter, as well as Venus. Thermal radiation from Saturn has also been measured. In this kind of measurement, however, there are considerable difficulties and as yet the work must be repeated many times before any final conclusions can be drawn.

In 1946 and 1947 A. N. Stewart in England first used radio equipment to study meteors. This team were the first group to give definite proof that the transient echoes from the "E" layer of the ionosphere were from meteors. Hey, Parsons and Stewart made the first radar measurements of meteor velocities.

An outstanding discovery in radio astronomy was made in 1951 by Ewen at Harvard University by the detection of the 1,420 megacycle hydrogen line. It was first suggested by Van de Hulst in Leiden University in 1944 that this radio spectral line might be detected in the galaxy. It is interesting to note that this investigation by Van de Hulst was stimulated by a paper published by Reber on "Cosmic Static". This hydrogen line is of considerable importance because a major fraction of the interstellar matter is made up of atomic hydrogen. It is not possible to detect this by optical means and we have therefore in this hydrogen line a new tool for the investigation of the galaxy and the universe.

An attempt was made in 1955 by Lilley and McClain in the United States Naval Research Laboratory to measure doppler shift by means of the hydrogen line. However, although they claimed a satisfactory correspondence between the radio and the optical measurements, this has never been repeated. At least one attempt was made by workers at Jodrell Bank who were unable to confirm the Americans' work.

In 1955 Burke and Franklin in America claimed that they had detected radio outbursts from the planet Jupiter at frequencies around 22 megacycles and again at 18 megacycles. Shain in Sydney examined past records back as far as 1950, and concluded that he had also recorded this emission. There are, however, a number of objections to these claims and during the past years some of the author's work has been devoted to this end. No confirmatory records have been obtained in this country so far, at the frequencies claimed, but it is significant that one American team found that these bursts of radiation appeared to coincide with outburst from the Sun during 1960.

Previously, in 1958 and 1959, the author had, from his own observations, concluded that the Sun played an important part in this mechanism. Several teams have been at work in America and a considerable amount of observational data has been collected, particularly by Carr and Barrow. An extensive observational programme has been planned by workers in America and Europe over the next few years. Many observations will be needed before any final conclusion can be drawn. It was claimed by Kraus in 1956 that he had detected radiation from Venus which he considered were similar to those alleged to come from Jupiter. However, when he attempted to repeat the observations later he came to the conclusion that he had been mistaken and that the radiations were merely disturbances of the ionosphere. It is certain that any extra-terrestrial radiations received below the frequency of 30 megacycles are subject to considerable modification by the ionosphere. Any radiations recorded, therefore, must be examined with great care for many disturbances can originate in the ionosphere itself. It is only recently that this view has become more widely accepted.

In the early 1950's only a handful of radio sources were known. Since then surveys have been made in England and Australia and now some thousands of radio sources have been accurately charted. There is still disagreement between the Australian survey and the Cambridge survey, but no doubt work which is proceeding now will eventually reconcile any discrepancies there may be.

In 1961 Professor Ryle and the Cambridge team completed their first survey of a new system of aerials called "aperture synthesis". The results of these observations were almost revolutionary for they indicated that the steady-state theory sponsored by Hoyle, Bondi and Lyttleton was not supported by these observations. The implications of all this is that the evolutionary theory of the universe is more tenable than that of the steady-state universe. At the moment the evidence is not sufficiently conclusive to say whether or not the steady-state theory must be discarded. Professor Hoyle and his colleagues are already offering an alternative explanation to that of Professor Ryle, and an attempt to check Professor Ryle's observations is now being made by Professor Sir Bernard Lovell with the Jodrell Bank 250-ft parabola.

Jodrell Bank will be remembered for its work on meteor astronomy and meteor physics. The measurement of star diameters included the notable feat of determining the diameter of Sirius, a star too small to be measured by Michelson's original interferometer method with an optical telescope. This work was carried out by Hanbury Brown and Twiss.

Work done in 1954 using the Moon echo technique gave us a great deal of information as to the surface structure of the Moon. The Moon echo technique also contributed to our knowledge of the Faraday rotation, which takes place when a radio wave passes through the ionosphere. The identification of radio sources by techniques which make use of the occultation of radio sources by the Moon itself are carried out. In Cambridge Professor Ryle together with members of his team, Dr. Hewish and Dr. Graham Smith and others, have added further lustre to the tradition of the Cavendish Laboratory. Here the first radio source measurements were made in 1946 using the new technique of the radio interferometer with consequent higher sensitivity and greatly increased resolving power. Cambridge provided the first accurate maps of the position of radio sources and the determination of their angular diameters. From these measurements the radio sources Cygnus A and Cassiopeia A were identified. Reference has already been made to other observations by the Cambridge group where the world's largest radio telescopes are situated at Lords Bridge.

The Mullard Radio Observatory, as it has been named in recognition of the large sum of money donated by Mullard Limited, affords Professor Ryle and his team opportunities for the exploration of the farthest reaches of the universe. A recent grant by the Department of Scientific and Industrial Research has made it possible to begin the construction of an even larger telescope for the purpose of extending the activities of the Mullard Radio Observatory. This will enable observations to be carried out up to a distance of 10,000 million light years; the distance at present regarded as the observable limit of the universe.

At Malvern Dr. J. S. Hey continues to pursue his investigations in the field of radio astronomy. With his new interferometer consisting of two 85-ft diameter steerable dishes, he plans a very extensive future programme of investigation. New information is expected to be obtained about conditions and the radio processes within the solar system itself. Research of this type is now very important for the applications of military and civil techniques involving missiles. This new installation was devised to be extremely accurate in positional measurements and provide a versatile system rather different from that in use at Cambridge or Jodrell Bank. The aerials have been designed to work down to wavelengths of 10 centimetres. The variable spacing arrangement that is possible with this instrument makes it specially suitable for certain types of radio and atmospheric research. Observations will be made of the effect of the upper atmosphere on radio waves passing through it and being reflected by free electrons in meteor trails.

To the Jodrell Bank installation is to be added still another giant telescope. This new design, based on the experience gained with the 250-ft telescope, will be used sometimes on its own and sometimes in conjunction with the original telescope. It has an elliptical shape with a major axis of 125 ft and a minor axis of 83 ft. The power gain of this new telescope will be at 10 centimetres as good as that of the 250-ft bowl at 30 centimetres, which is a very considerable improvement. Part of the programme will be devoted to lunar occultation of radio sources. A certain amount of its time will be devoted to investigation of the 21-centimetre hydrogen line and the observation of extra-galactic nebulæ.

In Australia, the Radio Physics Laboratory at Sydney is the largest in the southern hemisphere. Here Wild developed the radio spectrograph with which he first demonstrated the rapidly moving disturbances called "flares" which travel out through the solar corona at very high speeds from the surface of the Sun.

W. H. Christiansen was responsible for the design of the radio heliograph to examine the radiations from the Sun. This was made up of sixty-four paraboloidal reflectors in an interferometer which provided a resolution of 3 minutes of arc compared with the 30 minutes diameter of the Sun. With this instrument a complete map of the Sun's surface can be made.

The Mills Cross is another very large aerial invented by B. Y. Mills, and is a telescope of high resolution. The Australian radio star chart was compiled from measurements made with this instrument. There have been many active Australian workers, J. H. Piddington, O. B. Slee and the late C. A. Shain, to mention but three.

In America there are now some twenty radio observatories in operation and this new science is being pursued with great enthusiasm.

In the land of the birth of radio astronomy an American scientist speaking at the opening of the Owens Valley project, sponsored by the California Institute of Technology, said: "The project is one of several designed to elevate radio astronomy work in this country to that being done in Australia, Britain and the Netherlands. America has lagged after pioneering the new science." The Owens Valley project was an extensive one. Its main equipment consists of two 90-ft paraboloids each equatorially mounted and movable on tracks to form part of a large interferometer. Part of its observing programme is to attempt to identify some 1,500 or more discrete radio sources not yet correlated with visible objects. The range of operation is to be between 5 centimetres and 1 metre.

At Palo Alto, in California, R. M. Bracewell and his team at Stamford University have been working on a high resolution interferometer. This is designed chiefly for mapping the Sun's radiation and consists of sixteen 10-ft parabolic dishes set in a row 375 ft long and bisected at right angles by another row of sixteen; this is in the form of the Mills Cross. The resolving power of this interferometer is equal to that of a paraboloid 375 ft in diameter. It is possible with this installation to observe areas as small as 0·003 square degrees of the surface of the Sun.

Another extensive installation has been set up by the National Bureau of Standards at Boulder, Colorado. Here again solar emission is studied in the range 167 megacycles to 460 megacycles. Nearby at Table Maser two 40-ft radio telescopes are used for the study of scintillation of radio sources. Other paraboloids are planned for this installation in order to extend their activity.

The same American organization has set up aerials for the study of meteors by radar techniques at Havana in Illinois. There will be six aerials in all, five of them spaced on a straight line 35 miles long. It will be possible with this aerial to detect sufficient reflected energy from meteor trails to enable the paths and radiance of the meteors to be determined. These observations will augment the work being done at Jodrell Bank in England.

At Millstone Hill, at Westford, Mass, a large aerial is installed principally for the study of artificial satellites and rockets. This also can extend its activities into other fields.

These are but a few of the permanent installations and this work is being increased year by year as required.

In Europe, France has two radio observatories, mostly devoted to work on the Sun. Germany has an 80-ft paraboloid which is used for 21-centimetre work. At Leiden, under Professor J. H. Oort, there is an 85-ft parabolic reflector which is used for the 21-centimetre hydrogen line observations and other galactic studies. In the U.S.S.R. there are three radio observatories. Here again solar observations and 21-centimetre hydrogen line investigations are carried out. In Japan there are five radio observatories all working on the Sun. In Italy at the observatory at Arcetri there is a fully steerable radio telescope with helical aerials which continuously monitor the activity of the Sun.

This list is by no means exhaustive, but it is clear that there is still much room for expansion. The needs of satellite tracking and space navigation has called into operation many of the techniques of radio astronomy. It is certain that intensive work on apparatus will contribute both to the needs of communication and the investigation of outer space.

Herein, then, lies the opportunity for the amateur to make his contribution.

Chapter Three

BASIC ASTRONOMY

ONE OF THE first impressions gained by an observer who looks at the sky on a starlight night is that the brightest stars seem to form well-marked groups which constantly retain the same shape. One of the most conspicuous of these groups, or constellations as they are called, can be seen when looking towards the north. The group is called The Plough, or to give it its correct name Ursa Major, or the Great Bear; it is made up of several principal stars visible to the naked eye. The ancient astronomers ascribed to these conspicuous groups the names of great figures in mythology: some were considered to resemble men and some to resemble animals.

The Chinese, the Chaldeans and the Egyptians were the first people to divide the stars up in this peculiar manner. The earliest systematic grouping that we have on record is contained in a book called the *Almagest*. This book was compiled by the great astronomer Ptolemy about the year 140 A.D., and is a catalogue of stars divided into forty-eight groups, most of them being named after characters connected with the voyage of the ship Argo. In spite of the unscientific plan adopted by ancient astronomers in mapping the heavens the arbitrary division is, for the sake of convenience, still retained. The constellations may thus be considered analogous to the countries on the Earth with the stars as towns.

The majority of the bright stars possess proper names, such as, for example, Regulus, Aldebaran, Rigel, Betelgeux, Sirius, Altair and so on. Present-day astronomers use numbers and Greek letters; these are arranged so that the brightest stars in the group begin with the Greek letter alpha, thus Regulus is Alpha Leonis, that is to say, it is the brightest star in the constellation of Leo the Lion. Again, Aldebaran is designated Alpha Tauri, that is the brightest star in the constellation of Taurus the Bull. Astronomers adopt a scale as to the brilliancy of stars so that the brightness can be expressed by a special reference. This system works upon the basis that on the average the light received from one of the brightest stars in the heavens is a hundred times greater than that from a star which is only just visible to the naked eye. The difference of brightness lies between these two extremes and can be divided into a number of steps. The stars visible to the unaided eye are taken in an order of 6 degrees of brilliancy. The brightest are classified as stars of the first magnitude and the faintest naked-eye star is a star of the sixth magnitude.

When we look up into the sky on a clear night the stars appear to be countless in number, yet if we were to total them up there would be not more than 4,000 visible to the naked eye. We can classify these roughly with our brightness order of magnitude, thus there are about 14 stars of first magnitude, 48 of the second magnitude, 152 of the third magnitude, 313 of fourth magnitude, 854 of the fifth magnitude and about 2,000 of the sixth magnitude. Even a small telescope is sufficient to add very many more stars to our list than can be observed with the naked eye. With an aperture of less than 3 in. in diameter it is possible to see some 300,000 stars in the northern hemisphere.

We have already spoken of Ursa Major and its seven bright stars. Two of these are used as pointers to the Pole Star, and if a line is carried on through the Pole Star we shall come on to the opposite side of the sky to Cassiopeia. A full set of charts showing the positions of the various constellations will be found in Patrick Moore's book, *The Amateur Astronomer*. A list of the brightest stars visible is shown in Chart I.

Though all the stars are really moving through space they are so far removed from the Earth that their real motions are only detected after several years, even with very refined astronomical instruments.

So far as our observations are concerned, the stars may be regarded as fixed points on a celestial sphere. This is an imaginary sphere with the Earth at the centre and the stars, or any other celestial object, projected upon it. To locate the position of an object by two measurements, which are known as co-ordinates, it is necessary to have two lines, or planes, at right-angles to one another from which to measure. In the case of celestial objects the sensible horizon is the plane from

CHART I

THE BRIGHTEST STARS IN THE SKY

STAR	PROPER NAME	MAG.	SPECTRUM	DISTANCE IN LIGHT YEARS	LUMINOSITY SUN = 1
α Canis Majoris	Sirius	— 1·43	A1	8·6	26
α Argus	Canopus	— 0·73	F0	650	80,000
α Centauri	—	— 0·27	G2	4	1·1
α Boötis	Arcturus	— 0·06	K2	41	100
α Lyræ	Vega	0·04	A0	26	50
α Aurigæ	Capella	0·09	G0	47	150
β Orionis	Rigel	0·15	B8	540	18,000
α Canis Minoris	Procyon	0·37	F5	10	5
α Eridani	Achernar	0·53	B3	66	200
α Orionis	Betelgeux	var.	M2	190	1,200
β Centauri	Agena	0·66	B0	300	3,000
α Aquilæ	Altair	0·80	A7	16	9
α Tauri	Aldebaran	0·85	K5	57	90
α Crucis	Acrux	0·87	B0	230	1,000
α Scorpionis	Antares	0·98	M1	360	3,400
α Virginis	Spica	1·00	B1	230	1,500
α Piscis Australis	Fomalhaut	1·16	A3	24	13
β Geminorum	Pollux	1·16	K0	32	28
α Cygni	Deneb	1·26	A2	650	10,000
β Crucis	—	1·31	B0	200	850
α Leonis	Regulus	1·36	B7	56	70

which the angular height or altitude is measured. The line from which the azimuth is measured is the line which joins the north and south points of the horizon. When we have measured the altitude of a star we know that it is at a certain distance above the horizon and then, if in addition to this, we measure its azimuth, we are made aware that it is at a distance from the north and south line. From the diagram in Fig. 4 it will be understood that the position of a celestial object at any instant can be defined by the altitude and azimuth system of co-ordinates. For certain purposes this method of location of a celestial object is useful, but we must bear in mind that when bodies are moving there is a continual change of altitude and azimuth. It is not, therefore, convenient to use these co-ordinates for defining the absolute position of an object upon the celestial sphere, but it is a simple method of locating the position of a star which once found can then be followed by the telescope of the observer. If, however, we wished not only to locate the position of the star on the celestial sphere but also follow its apparent motion automatically without the attention of the observer, two other co-ordinates are used for its location. These are right ascension and declination.

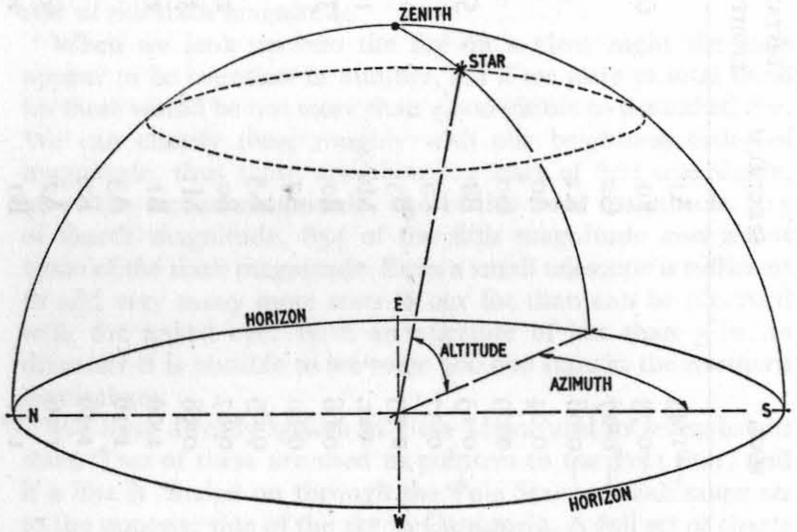


Fig. 4 Altitude and azimuth position of a star

Right ascension is analogous to the longitude on the Earth and declination is analogous with latitude. The celestial poles may be regarded either as the points in the celestial sphere directly above the poles of the Earth, or as points where the Earth's axis is produced and meets the celestial sphere, see Fig. 5. The celestial equator is a line drawn round the sphere half-way between the poles; circles passing through both poles thus cut the celestial equator at right angles.

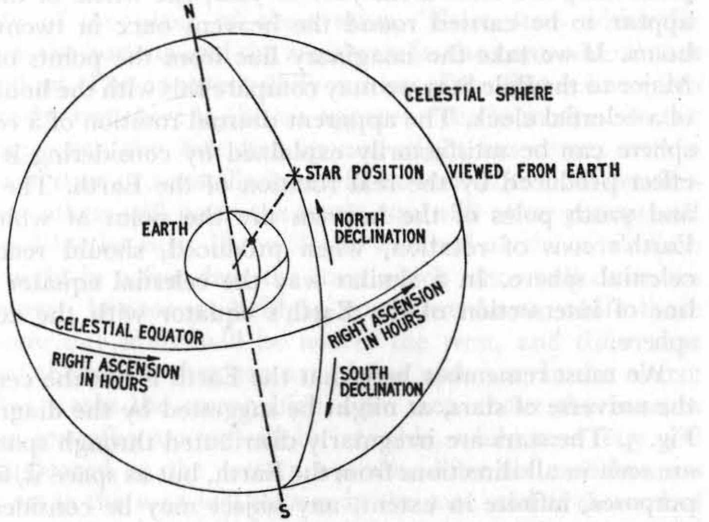


Fig. 5 Illustrating right ascension and declination

The celestial co-ordinate term "declination" may be defined as the angular distance north or south of the celestial equator. This measurement is made in degrees, minutes and seconds of arc.

On March 21 in each year the Sun is directly in front of one of the points of intersection known as the first point of Aries. This is the point from which right ascension is reckoned, just as terrestrial longitudes are measured from Greenwich. Right ascensions are generally reckoned from 0 hours to 24 hours of sidereal time from west to east, that is to say, in the opposite direction to the apparent diurnal movement of the heavens. The rotation of the Earth will now carry the telescope along with it and will sweep out a path equivalent to the right ascension and declination. By driving the telescope in the reverse direction at the equivalent speed of the rotation of the Earth we are able to keep the telescope fixed on the position of the star in space.

Though the configuration of stars observed in the sky remains practically the same from year to year, the whole of the stars appear to be carried round the heavens once in twenty-four hours. If we take the imaginary line from the points of Ursa Major to the Pole Star we may compare this with the hour hand of a celestial clock. The apparent diurnal rotation of a celestial sphere can be satisfactorily explained by considering it as an effect produced by the real rotation of the Earth. The north and south poles of the heavens are the point at which the Earth's axis of rotation, when produced, should reach the celestial sphere. In a similar way the celestial equator is the line of intersection of the Earth's equator with the celestial sphere.

We must remember here that the Earth is not the centre of the universe of stars, as might be suggested by the diagram in Fig. 5. The stars are irregularly distributed through space and are seen in all directions from the Earth, but as space is, for our purposes, infinite in extent, any object may be considered as the centre of the celestial sphere.

The north pole of the heavens is on the horizon of an observer situated at the equator. It appears higher and higher in the sky as we move northward until when we reach the north pole it is exactly overhead. There is a close connection between latitude and the angular distances of the celestial pole from the horizon; the latitude or angular distance from the equator or any spot on Earth is equal to the latitude of the celestial pole of that place, hence a rough determination of the latitude of a place may be made by noticing the altitude of the Pole Star above the horizon. From the north point of the horizon to the zenith is an angle of 90 degrees, and from the zenith to the south point is also 90 degrees. The whole angle from north to south, therefore, is 180 degrees.

The celestial equator is, of course, constant for any one latitude in the same way as the altitude of the celestial pole is constant for any one particular place. No definite line of stars marks the position of the equator, but its general direction can be determined by remembering the rule already stated as to its altitude.

On account of the Earth's rotation each star appears to describe a diurnal circle round the celestial pole, and the radius of the circle depends on the distance of the star from the Pole. The centre of the diurnal circle has an altitude equal to the latitude of the place of observation. These stars which lie between the horizon and the zenith are in the zone where stars rise and set. The apparent daily motion of the Sun is known to everybody, but there is also an apparent yearly motion. In the evening, when the Sun has set and night has come, certain groups of stars or constellations will be seen above the western horizon, others will be in the south, and still other groups will be found in the east. If the sky is watched at the same time every night in a few days the stars which were well down on the western horizon at twilight are lower down, while those that were due south will be nearer the west, and those that were on the eastern horizon are higher above it. In the course of a few weeks the stars which were seen above the western horizon just after sunset will be invisible, while new stars will have appeared on the eastern horizon. After six months stars which set in the west will appear in the east, and at the end of twelve months they will begin their circuit again. Near the equator where twilight lasts only for a short time the stars become visible very soon after the Sun has disappeared below the horizon. The constellations near the Sun at sunset are therefore more easily distinguished than in the middle latitudes. If they are noted night after night the apparent yearly path of the Sun among the stars can be found. This path is called the ecliptic. During the course of a year it cuts through the celestial equator at two points. One of these points we have already mentioned, that is the first point of Aries or the position from which right ascension is measured.

The other point that cuts the celestial equator is 180 degrees distant from it, or twelve hours right ascension.

The accumulation of observations showed early astronomers that the Sun was always accompanied by a particular star group at a particular time of the year. The Sun, therefore, keeps a fixed track, and it is this track which we call the "ecliptic".

Whatever part of the ecliptic the Sun occupies the stars in the neighbourhood are invisible at this point, for when an observer is turned towards that part of the heavens the Sun is above his horizon and the light of the stars is lost in sunlight glare.

Remembering that the position of any celestial object can be defined by means of the co-ordinates declination and right ascension, just as definitely as any point on the Earth can be defined by latitude and longitude, a method of determining exactly the apparent annual path of the Sun can be easily understood. All that it is necessary to do is determine the declination and right ascension of the Sun at a number of times throughout the year. We then plot the points thus obtained on the chart having lines representing declination and right ascension drawn upon it. The following table shows the right ascension and declination of the Sun at noon near the middle of every month.

TABLE 1

RIGHT ASCENSION AND DECLINATION OF THE SUN

DATE	RIGHT ASCENSION OF SUN AT NOON	DECLINATION OF SUN AT NOON	REMARKS
March 21	0h 1m	0° 0'	Spring Equinox
April 15	1h 36m	10° 3' N	
May 15	3h 31m	19° 3' N	
June 20	5h 58m	23° 27' N	Summer Solstice
July 15	7h 41m	21° 25' N	
August 15	9h 41m	13° 50' N	
September 23	12h 3m	0° 0'	Autumn Equinox
October 15	13h 23m	8° 49' S	
November 15	15h 25m	18° 42' S	
December 21	18h 1m	23° 27' S	Winter Solstice
January 15	19h 47m	21° 10' S	
February 15	21h 54m	12° 44' S	

In consequence of the apparent eastward motion of the Sun among the stars, different constellations are visible at different times of the year. The Sun's apparent path can be precisely traced through the year in the manner already described. Therefore, knowing the position of the Sun upon the celestial sphere at any time, the stars and constellations which form the background are also known for their fixed points upon the celestial sphere. These groups of stars, which lie in the neighbourhood of the Sun's track, are called the "Zodiacal Constellations" or signs of the Zodiac. The Zodiac itself is a zone extending completely round the heavens between the limits of 9 or 10 degrees north and south of the ecliptic. The positions of the signs of the Zodiac upon the celestial sphere are presented in Fig. 6.

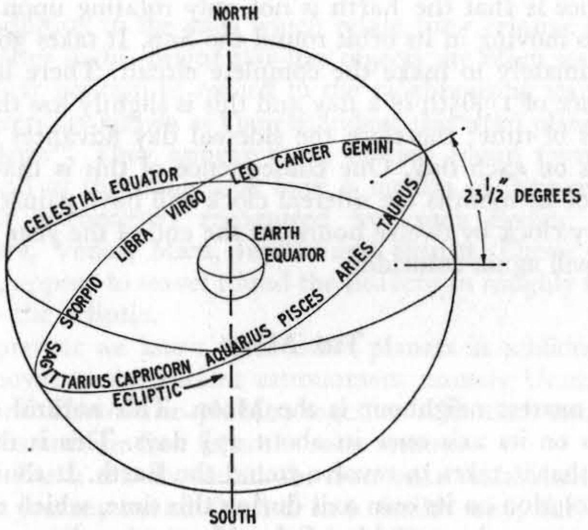


Fig. 6 Positions of constellation signs of the Zodiac

There is an old rhyme which enables us to remember the signs of the Zodiac in their order:

The Ram, the Bull, the Heavenly Twins;
and next the Crab the Lion shines;
the Virgin and the Scales,
Scorpion, Archer and He-goat;

the Man that bears the Watering Pot,
and the Fish with the Glittering Tails.

These are the names of the twelve divisions of the Zodiac.

All the apparent motions of the stars are due to the fact that the Earth is revolving on its axis. By noticing the interval between the passage of a star across the meridian on two successive nights we can tell how long it takes the Earth to rotate once on its axis. We find that there is an interval of approximately 23 hours 56 minutes; this time interval is called a "sidereal day" and is the true time taken by the Earth in making one revolution on its axis. In one sidereal hour the Earth turns one twenty-fourth part of a revolution, or through 15 degrees. We must here mention that there is a difference between ordinary time and sidereal time. The reason for this difference is that the Earth is not only rotating upon its axis but also moving in its orbit round the Sun. It takes 365¼ days approximately to make the complete circuit. There is thus a difference of 1/365th of a day and this is slightly less than four minutes of time; therefore the sidereal day advances by four minutes on each day. One consequence of this is that in the course of six months the sidereal clock will have gained on the ordinary clock by twelve hours. At the end of the year the two points will again coincide.

THE MOON

Our nearest neighbour is the Moon. This natural satellite revolves on its axis once in about 27⅓ days. This is the same period that it takes to revolve round the Earth. It thus makes one revolution on its own axis during this time, which explains why we see only one side of the Moon turned towards the Earth. A few days after the new moon it is often possible to see the Moon as a globe having the crescent of its illumination on the side that faces the Sun. This phenomenon is known as "old Moon and the young Moon's arms". The explanation of this appearance is comparatively simple. The Sun illuminates the Earth as well as the Moon and the Earth reflects part of the light thus received. This reflected light strikes the Moon and faintly illuminates the hemisphere of the Moon which is turned towards the Earth, and causes it to be dimly seen. We thus see it, as it were, by second-hand sunlight, or as it is termed "Earthshine".

One peculiarity that will be noticed is that the bright crescent of the Moon appears to be part of a body rather larger than that of the dark portion of it; this is the effect of irradiation, where a bright object appears larger than a dark one to the eye and its image tends to spread out.

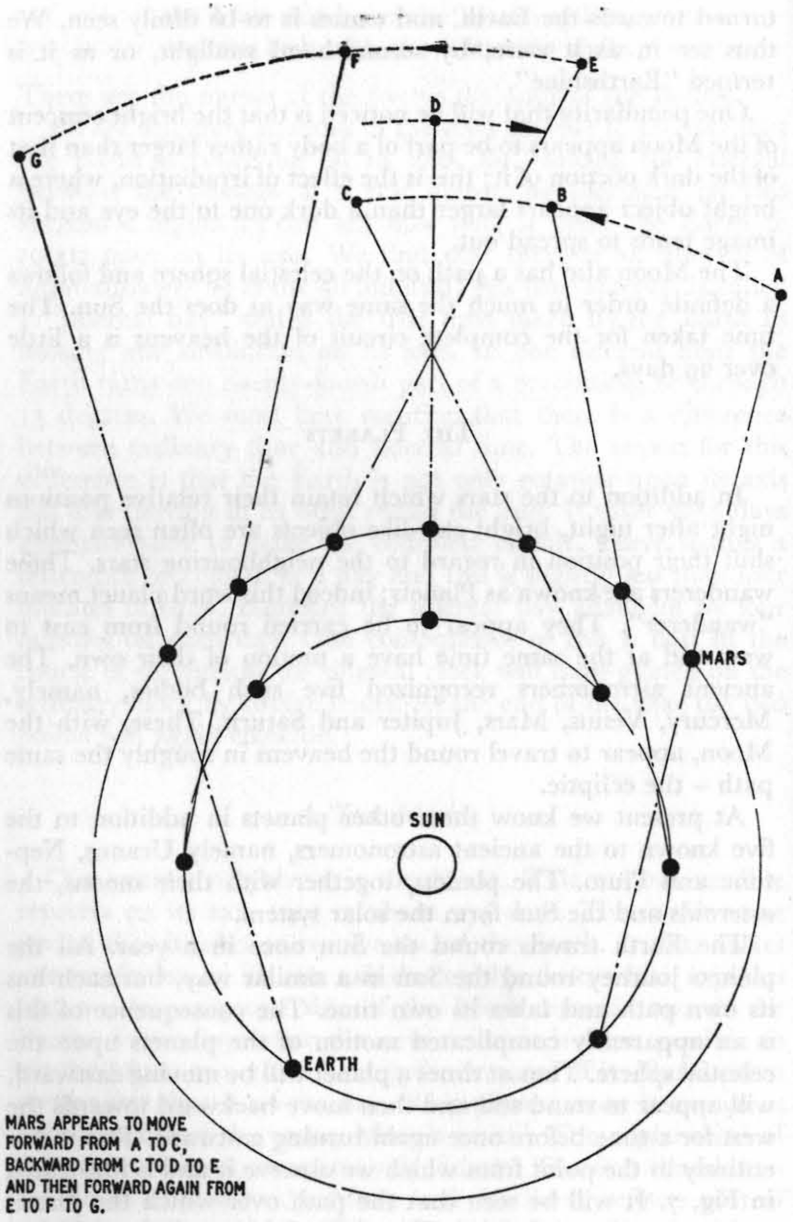
The Moon also has a path on the celestial sphere and follows a definite order in much the same way as does the Sun. The time taken for the complete circuit of the heavens is a little over 29 days.

THE PLANETS

In addition to the stars which retain their relative positions night after night, bright star-like objects are often seen which shift their position in regard to the neighbouring stars. These wanderers are known as Planets; indeed this word planet means "wanderer". They appear to be carried round from east to west and at the same time have a motion of their own. The ancient astronomers recognized five such bodies, namely, Mercury, Venus, Mars, Jupiter and Saturn. These, with the Moon, appear to travel round the heavens in roughly the same path – the ecliptic.

At present we know three other planets in addition to the five known to the ancient astronomers, namely Uranus, Neptune and Pluto. The planets, together with their moons, the asteroids and the Sun form the solar system.

The Earth travels round the Sun once in a year. All the planets journey round the Sun in a similar way, but each has its own path and takes its own time. The consequence of this is an apparently complicated motion of the planets upon the celestial sphere. Thus at times a planet will be moving eastward, will appear to stand still and then move backward towards the west for a time before once again turning eastward. This is due entirely to the point from which we observe it and is illustrated in Fig. 7. It will be seen that the path over which the planet appears to move is looped. The time of the revolution of Mars round the Sun, for example, is one year ten months. While the



MARS APPEARS TO MOVE
FORWARD FROM A TO C,
BACKWARD FROM C TO D TO E
AND THEN FORWARD AGAIN FROM
E TO F TO G.

Fig. 7 Retrograde movement of Mars as seen from Earth

Earth is moving from A to B in its orbit Mars is similarly moving from A to B in its own orbit, and the succeeding positions of the two bodies after equal intervals are shown by consecutive letters. The dotted lines passing through corresponding positions of the Earth and the planet under observation locate the apparent situation of the planet in space to a terrestrial observer. If these localities are joined by a curve, the loop path which is shown in the figure is obtained.

It should be mentioned that when a planet or any other member of the solar system is moving eastward among the stars, it is said to be in direct motion. Motion towards the west is described as retrograde. The planet is said to be in conjunction when it is in the same direction as the Sun, and so comes on the meridian at the same time. It is in inferior conjunction if it lies between the Earth and the Sun, and in superior conjunction if on the remote side of the Sun. When the Earth is not only in the same straight line with the planet and the Sun, but between the two, the planet crosses the meridian at midnight, and in these circumstances is said to be in opposition. The apparent angular distance between the position of the Sun and that of the planet as observed from the Earth is known as the elongation of the planet. When the elongation is 90 degrees the planet is in quadrature. It will at once be seen that when the elongation is 0 degrees the planet is in conjunction, and when it is 180 degrees the planet is in opposition. Only planets outside the Earth's orbit can be in opposition; these are known as "Superior Planets".

VENUS

What is usually known as the Morning and Evening Star is the planet Venus, the second planet in order of distance from the Sun. The planet Mercury also appears as a morning and evening star, but it is much fainter than Venus and can only be seen just above the Sun at sunrise or sunset. Many astronomers have never seen this planet. When Venus is visible in the sky it is by far the brightest object in the heavens. If it is watched day by day, after it has just appeared, as an evening star, its position seems to change steadily, becoming more and more eastward. After a time this eastward movement ceases and the planet remains stationary for a day or two; then a westward progression is observed, the planet getting closer and closer to the Sun at sunset and is finally lost in the twilight glow. After a while it appears on the other side of the Sun and is seen as a morning star; it moves to a maximum of about 48 degrees away from the Sun and then goes back again. Venus thus appears to oscillate to and fro between 48 degrees west and 48 degrees east of the Sun. Of course, the planet does not really swing to and fro in this way, but only seems to do so on account of its own motion and the Earth's motion round the Sun. Its path with reference to the stars is like that of the other planets. The appearances of Venus before sunrise and after sunset, which have given rise to the popular expression "Morning and Evening star", are easily understood.

Venus is one of the planets about which we know extremely little; no one has ever seen its surface, so it is assumed to be completely surrounded by cloud. It is thought that this cloud covering may consist largely of carbon dioxide. Many attempts have been made to discover its period of rotation, but though there are some eighty-four different estimates we do not know which of these is correct. The latest measurements which have been made suggest that it might be a matter of a few hours. It will be necessary to wait until more refined measurements can be made before we can accept any one of these periods which vary from a few hours to more than 200 days. It is not surprising therefore that one of the major projects in space research is to send an observational vehicle as close to Venus as possible.

Venus is approximately the same size as our Earth, it being 7,680 miles in diameter; it is 67½ million miles from the Sun and takes some eight months to complete its orbit round the Sun.

MERCURY

Mercury is situated some 36 million miles from the Sun and is some 3,000 miles in diameter. It takes about three months to complete its orbit round the Sun. It appears always to turn one face towards the Sun, just as does the Moon towards the Earth. It would therefore seem that its rotational period is the

same as that of its orbital period, that is to say, it revolves once on its axis during its orbit round the Sun. It must for this reason be extremely hot on its sunlit side, and the surface temperature there has been calculated to be 340 degrees Centigrade at maximum.

The superior planets, or those which are farther away from the Sun than the Earth, are Mars, Jupiter, Saturn, Uranus, Neptune and Pluto. Three of these, Mars, Jupiter and Saturn are conspicuously visible to the naked eye.

MARS

Mars is much smaller than the Earth, being some 4,200 miles in diameter. Its rotational period is slightly longer than our own, some forty minutes in fact; but it takes nearly two years to complete its orbit round the Sun, from which it is distant about 142 million miles.

In 1877 Schiaparelli announced that he had discovered a pattern of dark channels on the surface of Mars. This was translated into English as being canals, and at one time these canals were taken to be evidence of life on the planet. It was suggested that these canals were made for the purpose of irrigating the surface of the planet. We do not now believe this. That these changes in surface markings take place, there is no doubt; they have been photographed and observed over many years. As the seasons shift from one hemisphere to another the appearance of Mars changes in a more or less regular way and we are able to observe the melting of the white polar caps during the Martian summer. Other formations that are visible on parts of the disk show seasonal changes in intensity and colour. Recent investigations suggest that it is possible that there is some form of vegetation, perhaps a very low type of vegetation in the form of moss or lichen. Kuiper, an American astronomer, has shown that the polar caps consist of frozen water and snow; the reddish yellow parts of the planet are most probably a kind of desert. The atmosphere on Mars is very much thinner than that of the Earth, but because of the evaporation of the polar caps it is assumed there must be a certain amount of water vapour present. This, however, has never been detected spectroscopically; neither, in fact, has oxygen been detected in this atmosphere. It is possible that there is nitrogen and argon. Any form of life, therefore, that might be present would be rather different from that which could exist in the kind of atmosphere that we have on Earth. Mars, being further away from the Sun than the Earth, has a much lower surface temperature. At the poles this may rise to about −15° C during the summer; the dark zones which may possibly represent vegetation can be as hot as +50° C. The desert regions of the planet are probably cooler. At sunset the temperature would normally be about 0° C and fall to something like −20° C during the night. By our standards, therefore, Mars would have a very severe climate.

Mars has two satellites, Phobos and Deimos. Phobos is some 6,000 miles distant from the centre of the planet and Deimos is something less than 15,000 miles distant from it. The diameters of these satellites are not known accurately since they are both very small. It has been suggested that Deimos is perhaps 5 miles in diameter and Phobos 10 miles in diameter. Phobos has an orbital period of 7½ hours, which means that it rises and sets more than once during the course of a Martian day.

JUPITER

The next planet is the giant Jupiter. It is some 483 million miles from the Sun and completes its solar orbit in a little under twelve years. Its diameter is approximately 88,000 miles. Since it is mainly gaseous, different parts of the planet have different rotation periods; the mean rotational period is about 9 hours 50 minutes. Because of its high rotational speed there is pronounced centrifugal flattening, the difference between the polar diameter and the equatorial diameter being some 6,000 miles. Jupiter's atmosphere appears to be very dense, and ammonia and methane have been detected. The temperature of the visible part of the planet is around −130° C. It is an extremely interesting planet because it is marked by a number of well-defined bands which are known as belts. It also has certain peculiarities which take place in the configuration of these belts which tend to change in shape and density. There is a large area known as "the Red Spot" which appears regularly.

Jupiter has twelve satellites, the nearest being some 112,000 miles distant from the planet and the furthest some 15 million miles. The details of these are shown in Chart 2.

CHART 2

JUPITER'S SATELLITES

NO.	NAME	DIAMETER MILES	DISTANCE (10^3 MILES)	REVOLUTION PERIOD
V	Amalthea	100?	113	0d 11h 50m
I	Io	2,310	262	1d 18h 27m
II	Europa	1,750	417	3d 14h 13m
III	Ganymede	3,200	666	7d 3h 42m
IV	Callisto	3,220	1,170	16d 16h 32m
VI	(Hestia)	100?	7,120	250d 16h
VII	(Hera)	30?	7,290	259d 16h
X	(Demeter)	25?	7,300	260d 12h
XII	(Adrastea)	15?	13,000	625d
XI	(Pan)	20?	14,000	692d
VIII	(Poseidon)	35?	14,600	739d
IX	(Hades)	20?	14,700	758d

Jupiter is the subject of constant observation by astronomers, particularly amateurs.

SATURN

The next planet in line, as we move outwards from the Sun, is Saturn. This is nearly twice as far away as Jupiter, being some 886 million miles distant from the Sun. It takes nearly thirty years to complete its orbit round the Sun, and its rotational period is again a high one, being some 10¼ hours. It also appears in the telescope as a somewhat flattened disk; its diameter is some 67,200 miles, measured along the polar axis, and some 75,000 miles at the equator. Its atmosphere appears to be similar to that of Jupiter; it is very dense and contains a large quantity of ammonia and methane. The temperature is even lower than that of Jupiter. Saturn, however, is unique among the planets in that it has a ring system; these rings which are approximately 169,000 miles in extent and about 12 miles thick form a beautiful sight when the planet is at an angle to us on Earth. The ring system is divided into three more or less distinct units; the middle one is brightest and they must consist of many small particles. Kuiper has suggested that the particles themselves may be composed of frozen water, somewhat as the form of hoarfrost. Previous suggestions and those most generally accepted are that a satellite of Saturn approached too close to the planet and disintegrated, and that the rings are made up of tiny particles of this debris. Saturn has at present nine known satellites, which are listed in Chart 3.

CHART 3

SATURN'S SATELLITES

NO.	NAME	DIAMETER MILES	DISTANCE (10^{3} MILES)	REVOLUTION PERIOD
I	Mimas	300	113	22h 37m
II	Enceladus	400	149	1d 8h 53m
III	Tethys	600	183	1d 21h 18m
IV	Dione	600	235	2d 17h 41m
V	Rhea	850	328	4d 12h 25m
VI	Titan	3,000	760	15d 22h 41m
VII	Hyperion	250	920	21d 6h 38m
VIII	Iapetus	750	2,200	79d 7h 56m
IX	Phœbe	150	8,050	*550d 10h 50m

* denotes retrograde motion.

URANUS

The remaining three planets are all telescopic objects. Uranus, the next one out from Jupiter, was discovered accidentally by Herschel, who at first wanted to name it after

George III. It has a diameter of some 29,300 miles and is nearly 1,800 million miles from the Sun. It takes eighty-four years to complete its orbit, but like Saturn has a very high period of rotation – some 10¾ hours. It has five satellites whose orbits are all in one plane. Uranus, however, has a peculiarity for the plane of its satellites is almost at right-angles to the ecliptic, that is to say, the polar axis of the planet lies almost parallel with the ecliptic. This results in our appearing to see the satellites of Uranus moving in circles around the planet when the pole faces us, almost as though we were looking straight down on it.

NEPTUNE AND PLUTO

Of the last two of the superior planets, Neptune and Pluto, very little is known. Neptune is at a distance of 2,800 million miles from the Sun; it completes its orbit in 165 years and the rotational period is believed to be fifteen to sixteen hours. Its diameter is some 28,000 miles. It was discovered in 1846 as the result of mathematical predictions which were based on the irregularities of the motions of Uranus calculated by Adams and Le Verrier. It has two satellites, Triton and Nereid.

Pluto was discovered in 1931 by Tombaugh. Except for the details of its orbit we know practically nothing about it; the diameter is some 3,600 miles and no satellite has yet been discovered. It is some 3,700 million miles from the Sun and takes nearly 250 years to complete its orbit.

THE ASTEROIDS

The asteroids, or minor planets, were unknown in ancient times and indeed for nearly two centuries after the invention of the telescope. The reason for this was that they were so small and very faint that they were almost impossible to distinguish among the myriads of stars in the sky background.

The first asteroid was discovered by chance by Piazzi at Palermo in 1801. Piazzi noticed that a star-like object, which he had been observing, appeared to move against the background. This was, in fact, the asteroid to which he gave the name of Ceres. Soon after the discovery of Ceres, Olbers discovered the second minor planet which was called Pallas, and so in 1802 a second minor planet was added. In 1804 Harding found another, which was named Juno. In 1807 Olbers again discovered another, and this he called Vesta. Vesta is the brightest of the asteroids and at favourable positions is just visible to the naked eye.

It was not until 1845 that the further members of the family of asteroids were discovered. Hencke found the fifth and then a sixth two years later.

New interest was immediately aroused in the subject of asteroids and new ones were added to the list every year. Only patient and diligent searching revealed these very small bodies revolving in their orbits round the Sun. Goldsmith discovered fourteen between the years 1852 and 1861. By 1868 the number of asteroids which had been discovered amounted to 100 and by 1890 this number had risen to 300. Up until this time the procedure adopted by astronomers was to examine the sky through the telescope and compare it with a star chart or atlas. If one star was observed which was not shown on the chart it would be observed for several consecutive days to detect any movement which would then show that it was an asteroid or a planet. This kind of observation required very great patience and skill.

It was at about this time that observational methods began to change. The advent of the photographic method reduced the labour required in watching and also increased the accuracy of the discoveries. A photograph gives a reliable map of the whole portion of the sky at which the telescope is pointed. The results can then be examined at leisure and should there be an asteroid, or a stranger, on the chart it gives a slightly different image from that of a star. Since the telescope moves at the same rate as the stars, an asteroid or planet which is moving in relation to the star background will show up in different positions after successive exposures or it may leave a trailing image on the plate. The length of the image depends upon the length of the exposure and the speed of the asteroid's apparent motion.

At the present time some 2,000 of these tiny planetary bodies have been discovered. Though their orbits should be, according to Bode's Law, confined to a particular area, we find, in fact, that they are widely scattered, some of their orbits extending out in between that of Jupiter and Saturn. They do not all revolve in the same general plane as the other planets, their orbits being inclined sometimes as much as 30 degrees or 40 degrees to the plane of the solar system.

The size of these minor planets can be appreciated from the values which follow: Ceres is about 427 miles; Pallas about 280 miles; Vesta 240 miles and Juno 150 miles. The lesser asteroids are irregular bodies and not spheres as the rest of the planets. The smallest of these asteroids that has been detected can be no more than one or two miles across.

We have a great deal to learn about the small particles in our solar system, and it will still require many years of work with refined techniques to learn all the many things that we wish to know about them. We do not know whether they carry an atmosphere, though the probability is that they do not for they are so small in mass that it is unlikely that any of them could have retained an atmosphere. We do not at the moment know how they originated; one hypothesis suggests that they were formed from material when the solar system was itself formed, but could not condense into one large planet owing to the disturbing influences of the other planets that were being formed. The second hypothesis is that they are the broken fragments of a planet which was disrupted either through collision or becoming disturbed by being too close to another, or again, by rotating so rapidly that it burst apart. Some support for this latter theory is given by the fact that the asteroids are of irregular shape, the kind of shape that would be expected from an explosion of this kind.

The total material in the minor planets even if brought together in one planet would only be very small. The total volume would not exceed 1/1000th part of that of the Earth.

COMETS AND METEORS

There are two other types of bodies which we should mention in this chapter on basic astronomy. The first of these are comets. These are wanderers that come to us from the depths of space. The origin and precise nature of comets we do not yet fully understand; they obey the same laws as the other members of the solar system, sometimes moving in highly elliptical orbits which are frequently inclined at high angles to the orbits of the planets. They may move either direct or retrograde. A comet may appear suddenly in any direction from the depths of space and pass round the Sun and disappear in quite another part of the sky.

In 1957 two new bright comets appeared, Arend-Roland and Mrkos.

During the period that comets are visible they sometimes change considerably in appearance; some of this is due to the direction from which they approach us and some are real changes in the comet itself which are caused by the Sun. This is particularly the case of the tail, which, by the pressure of sunlight, is always pushed in the opposite direction from the Sun. Thus, as a comet goes round the Sun its tail tends to change position.

Cometary orbits may be of three different kinds, they may be elliptical, parabolic or hypobolic. Comets with open orbits, of course, are once seen and they never again appear.

There have been a number of famous comets and these will be dealt with in greater detail in the chapter on cometary and meteor measurements.

The other class of bodies that we must mention is that of the meteors and meteorites. These are smaller wanderers which belong to the solar system itself. They consist of small particles which become visible when they burn up in the atmosphere. Occasionally there are excessively large bodies which strike the Earth, causing widespread damage. Some of these meteors are made of rock and some contain a considerable amount of basic metals. When one of these small particles passes through the atmosphere it burns up due to the friction with the atmosphere itself, and these may often be seen at heights of 100 miles or more. The average height of such an appearance is usually 70 miles above the Earth. The length of the path through the atmosphere would be something of the order of 50 to 100 miles. Fainter meteors are somewhat more restricted in their path lengths, which may be only as much as 35 to 40 miles. Some are especially brilliant and these meteors are commonly known as bollides or fire balls. They are not mere moving points of light but balls of fire whose size is sometimes equal to that of the diameter of the Moon as seen with the naked eye; sometimes they are white, sometimes red, yellow, green or blue. They are so brilliant that in a number of instances their appearances have turned night into day; they leave an ionized trail behind them testifying to the amount of energy which has been given up during their progress through the atmosphere. They would have speeds of the order of 25 miles per second.

There are three classes of meteors: the naked eye meteors, whose numbers are more or less accurately determined and the total number which is visible from the Earth's surface is something of the order of one hundred million per day. Secondly, telescopic meteors, which by using a new technique revealed up to one thousand million a day, and lastly those detected by radar methods, about which more will be said later in the section on comets and meteors. The total mass of material which is discharged into the Earth's atmosphere or falling upon the Earth amounts to some 100 tons daily. Meteors tend to appear in close groups, or showers, and at certain times of the year there are very definite appearances of these meteors. Recent discoveries have suggested that the ingredients of life are carried by meteorites.

The existence of life on Earth depends upon particular temperature conditions, conditions in which large molecules can exist, and also, of course, the presence of water vapour and oxygen. Life, therefore, as we know it is not possible on any of the planets. No terrestrial life could exist on a planet like Mercury whose temperature on the one side is very high and on the other side very cold. Mars, as we have seen from the conditions of its atmosphere, is hardly favourable to support life; though the extremes of temperature are not as great as on other planets, the atmosphere is thin and there would not appear to be enough water, in any case, available on the surface. Venus is probably the most favourable planet to support life, though since we know so little about it, it is difficult to form an opinion as to the conditions that may exist upon its surface. The presence of carbon dioxide in the atmosphere, together with other gases, may mean that there is perhaps some form of vegetation. It does not seem logical that the Earth should be the only place on which life could exist. Our Sun is a "G"-type star, one of many thousands of second-class stars. Our whole solar system when viewed from quite a moderate distance would hardly be visible. No telescope that we know to exist at present could show planets similar to those of our solar system when viewed from the distance of our nearest star. The gravitational disturbances due to Jupiter, the largest of the planets, cannot disturb the motion of the Sun by any detectable extent. Our solar system, therefore, is but a small part in the universe; it is indeed but a small part in our own little island in the universe, this island which we call the Galaxy.

THE GALAXY

Our galaxy contains myriads of stars. It is a spiral system of large diameter and thin section. The distances involved are so great that we cannot use ordinary terrestrial terms such as miles, which we have used in describing the distances of the planets. Instead we use the unit known as the "light year". This is the distance that light will travel at its speed of 186,000 miles a second in the course of a year. Our island system is some 100,000 light years across, and our solar system is situated at a point some 25,000 light years from the centre. When we look at the centre of our system we see it as it was 25,000 years ago. The thickness of the system is of the order of 20,000 light years, and this is the reason for the appearance of the bright band across the sky which we know as the Milky Way. It takes on this appearance because we are looking towards the edge of our island universe; when we turn our attention to the sky at right angles to this we find that the stars appear to be more scattered. Viewed from outside it would appear as an irregular and rather thin disk characterized by a number of spiral arms. This is one of the commonest types of the galaxies.

Contained in the galaxy are stars in many stages of evolution; in our own galaxy, for example, we have in the Sword of Orion a diffused nebulosity which contains many stars which are extremely hot, and these are thought to be young stars coming into being. Within the galaxy itself there are many radio sources and these young stars, surrounded as they are by hydrogen gas heated to very high temperatures, are prolific sources of radiation. Sometimes stars explode and these are given the name of novæ or "new stars". One such, known as the Crab Nebula,

is a very intense source of radiation. These new stars suddenly come into being, something in their make-up suddenly causing them to become intensely bright, flaring up sometimes so brilliantly that they can be seen during the day, and then they gradually fade away until finally they are seen no more or become very thin.

Outside our own galaxy there are many millions of other galaxies; some of these in an early state of development and some which are more highly developed and have the appearance of spirals similar to our own. We are fortunate that we have one as a reasonably near neighbour, at a distance of some 2,000 million light years; this is the great Nebula in Andromeda. It is visible to the naked eye, but in a small telescope appears as a diffused area. Photographs taken with the 200-in. Mount Palomar telescope reveal it to have a spiral structure very similar to our own. We are thus able at first-hand to study a galaxy which must be in many respects identical with our own. We have a further confirmation of this, for in radio astronomical measurements our own galaxy appears to be almost spherical; when we plot the size of the Andromeda Nebula with a radio telescope we find that this also tends to be spherical. We are thus able from our own position in the solar system to examine our galaxy from the inside, while we are able to observe a similar type of galaxy by looking from the outside at the Andromeda Nebula.

We shall be dealing with galaxies and nebular peculiarities in later chapters in this book when we deal with the application of radio astronomy to the study of these distant island universes.

Chapter Four

RADIATIONS FROM THE SUN

To THE INHABITANTS of the Earth the most important star in the sky is the Sun. As stars go it is of comparative insignificance, being one of the many millions in the galaxy which are designated "G"-type stars. It has, however, a peculiarity so far as we are concerned that is of great significance. This peculiarity, which is of great importance to the radio astronomer, is that the Sun is a variable source of radiation.

We know a great deal about the Sun for it is at one and the same time a vast power-house consuming itself in the production of radiations, and a laboratory in which changes are taking place under conditions we could not possibly reproduce on Earth. For example, different parts of the Sun have different conditions. We are able to examine the effect of electrons which spiral round the lines of magnetic force under conditions where the gas is far more rarefied than we could possibly produce on Earth. In many of these studies the very existence of these electrons is only known because of our ability to detect radio waves they emit. We are able to study nuclear processes and changes over vast ranges of temperature; some of these temperatures could only be achieved on Earth by the explosion of hydrogen bombs.

This vast sphere of highly compressed gas, some 864,000 miles in diameter, is visible to us as a bright disk. This disk we call the photosphere. It has a well-defined edge because there is a sharp change of temperature-level between that and the surrounding gas of the Sun's atmosphere. The temperature of this cool surface is some 6,000° Kelvin. We use the term "degrees Kelvin" here to designate absolute temperature; the normal Centigrade scale of temperature begins at 0 degrees, this is the freezing point of water. In order to deal with conditions existing in the universe we need to have a measure of temperature below this level. Absolute zero is therefore at a point below 0° Centigrade; in fact, it is -273°. In this range of temperature,

from -273° to 0°, there are many gases which liquefy, for example, helium, hydrogen and oxygen. One of the constituents of the fuel used in rockets is that of liquid oxygen. At these temperatures oxygen can be carried in containers like water. Many peculiar effects are noted at these temperatures; some materials attain a condition known as super-conductivity, and under such conditions an electric current can be started in a ring of metal and will thereafter continue to circulate without added energy, while the temperature is maintained at this low level. Liquid helium is also used for the operation of special types of radio receivers.

The temperature at the centre of the Sun is, however, enormously high. From the size and mass of the Sun its central temperature is deduced to be 15-20,000,000° Kelvin. The density at the centre of the Sun is so great that it is a hundred times greater than that of water, yet under these conditions of terrific pressure it still remains a gas. Such conditions are exactly those for the production of energy by atomic fusion. About half the mass of the Sun is composed of hydrogen, and of what remains a large proportion is composed of helium. This enormous reactor in the sky radiates energy by convection from its central core to the surface of the photosphere. From the surface it escapes as radiation in all patterns of the spectrum. As a result of the nuclear processes which are going on in the interior of the Sun, it consumes about four million tons of its mass each second. This activity extends to a diameter of about 180,000 miles. The energy is conveyed from this central zone by the radiative transfer of photons. These move from atom to atom on their journey outward. Outside this radiating core is a convection zone which extends to within a few thousand miles of the photosphere, and from here the energy is conveyed by the upward motion of the heated material through a very steep temperature gradient; the actual thermal conduction is negligible.

Next to the photosphere is the chromosphere, and it extends outwards for a distance of about 12,000 miles. It is slightly pinkish in colour and it is from this that its name is derived, which is chromosphere or coloursphere. There is a layer about 600 miles deep at the lower levels of the chromosphere known as the "reversing layer". In this region there are many atoms which have lost only one electron; in this state they are said to be ionized. Under such conditions an atom can absorb visible light. There is continual activity, therefore, in this area by electrons jumping from one level to another in the atom; this results in the production of intense ultra-violet radiation which eventually reaches the upper atmosphere of the Earth and contributes to the production of the ionosphere. Intense energy welling up from the interior of the Sun is absorbed principally in the chromosphere. The temperature varies throughout the chromosphere and there seems to be three distinct levels: one which rises to 10,000° Kelvin and thence up to about 50,000° Kelvin in the middle levels, and then another tremendous rise to nearly 1,000,000° Kelvin when we reach the outermost layer at the beginning of the corona.

If we project an image of the Sun on to a screen it is found that the outer edge of the photosphere is darkened, and this condition is known as "limb-darkening". The reason for this is that when we look directly at the centre of the disk we do in fact see *into* the interior of the Sun. When we look towards the edge we see through the much lower temperature level and it, therefore, appears darker. When examined in monochromatic light the surface of the Sun shows very considerable granulation. This gives the appearance of small white grains rather like rice sprinkled at random on the background. These small areas of the order of 700 miles in diameter have an average life of about three minutes. It is probable that this granulation is the visible evidence of the activity of convection from below the photosphere; the bright areas are probably the tops of hot rising columns of gas which bring up energy from the interior. The difference in brightness between the darker and bright areas suggests there is a temperature difference of the order of 200° Kelvin. Great geyser-like projections are continually emerging up into the chromosphere from the surface of the Sun. These are known as spicules and pour into the chromosphere photons that eventually are able to escape into outer space. The excited state of the chromosphere also produces a number of absorption lines known as Fraunhöfer lines. These lines were originally observed by a scientist named Wollaston; later in 1816 the young German scientist, Fraunhöfer, made a special study of them. These Fraunhöfer lines are valuable

as they enable us to determine the basic elements making up bodies such as the Sun. Prominent among these are calcium, strontium, hydrogen, helium, iron and titanium.

Beyond the chromosphere we have the corona. Until the advent of radio astronomy we were only able to see this corona at the time of total eclipse, or by the use of a special instrument called a "coronagraph", which was designed by a French astronomer, Lyot. It enables conditions similar to that of the eclipse to be produced artificially. There are, however, severe limitations to its use for it can only be put into effective operation when there is the clearest possible atmosphere and after extreme precautions have been taken to exclude dust from the inside of the coronagraph itself.

To the radio astronomer the corona extends out to distances greater than ten times the diameter of the Sun. The corona visible to the optical astronomer extends very much less into space. The corona varies continually because it is agitated from below due to the activity from inside the Sun itself and the chromosphere. It sends out vast streams in all directions; there are times when it is in greater activity than at others. It has been observed that at the time of sunspot minimum the rays and plumes have their greatest extension round the equator of the Sun, while at the period of sunspot maximum the rays are much shorter and more uniformly distributed. There are, however, always well-defined areas near the poles of the Sun which are clear. The streamers from the corona follow the Sun's magnetic field when the Sun is in the condition that we know as "quiet".

There is a regular cycle of activity which varies between 10 and 11·2 years. The surface of the Sun is often marked with dark spots which increase in number as we reach a period of maximum activity. These are sometimes quite small and scarcely visible on the surface of the photosphere, while at other times they are so large in extent that they may be in chains as much as 100,000 miles long.

These sunspots are the visible signs of enormous areas of turbulence taking place within the Sun itself. They appear usually in a well-defined belt around the equator of the Sun; at the time of minimum activity they can extend high to the north and the south of the equator. The progress of these spots across the face of the Sun can be followed, and on some occasions they are seen for a full revolution of the Sun – some 27½ days. Sometimes a small number of spots coalesce and striking effects are noticed occasionally with them. Bright bands of light appear and at these times there is usually considerable radiation in the radio part of the spectrum. A large crop of sunspots does not necessarily mean that there will be a high level of radiation so far as the radio astronomer is concerned. These radiations are due to turbulences set up through the chromosphere into the corona. These flares, as they are called, are terrific outbursts of gas and sometimes they appear above sunspots. The reason for this is thought to be because of the intense magnetic fields which are set up in these regions. Under such conditions great spouts of gas can be ejected from the Sun complete with their own magnetic field. They shoot out together with particles into inter-planetary space, occasionally pouring into our upper atmosphere and causing intense changes and modifications to it. Sometimes these flares reach out only so far and then fall back again to the surface of the Sun.

Over the surface of the Sun there appear from time to time prominences. These are vast clouds of hot and very tenuous gases which attain extremely high temperatures. These prominences bear a relation to the sunspot cycle. About two years before maximum these prominences appear to be larger and more numerous in the higher latitudes of the Sun. Prominences which appear in both high and low latitudes are called "quiescent"; these are often of great length and rise to enormous heights. The average heights may be anywhere between 50,000 miles to 250,000 miles high, though in exceptional cases they have even been observed at heights of more than 500,000 miles. Some of them remain for many weeks, apparently quite permanent, and then suddenly blow up and disappear within a few hours. During this time of blowing up, these ragged flames are ejected with speeds of from 50 miles to 250 miles per second. Their composition is similar to that of the chromosphere, often with the addition of other metallic vapours. They are sometimes subject to rapid transformations and disintegrate occasionally into separate components which vanish one by one while they still continue to rise from the chromosphere. By the use of the coronagraph, records can be made of the movement

of these prominences; usually they are more frequently noticeable between latitudes of 20 degrees and 40 degrees north and south. These zones move towards the equator during the course of the sunspot cycle, dying out at sunspot minimum and then appearing again in the higher latitudes at the beginning of each new cycle.

We have already mentioned the three levels of the chromosphere. The elements of calcium and hydrogen in the form of clouds are distributed through the layers. These bright clouds have been named "flocculi". The flocculi most usually take the form of patches or clouds extending over vast areas and they are often found to develop above the faculæ associated with sunspots. The incidence of flocculi then, like the sunspots, varies with the eleven-year cycle.

There is, however, a regular network of small flocculi which is a more or less permanent feature of the chromosphere. This is particularly the case of the lower levels of the chromosphere, but it also applies to the upper levels. In these levels hydrogen is predominant, and both hydrogen and calcium are distributed in the immense clouds which completely obscure the spots. The clouds of hydrogen assume a very characteristic whirlpool appearance and for this reason they are often referred to as "solar vortices".

Another type of prominence is called "eruptive". These appear suddenly and last sometimes only a few seconds. They attain enormous speeds of up to 100 miles a second. Some of these slow up as they reach the regions of 20–30,000 miles above the chromosphere; they are then pulled back into the Sun with increasing velocity. Such a display of activity may last for twenty minutes or more; others are so violent that they go right out into space without any visible sign of return. They give rise to short-lived radio bursts of radio waves; these begin at the high frequencies and as they die away the radiation falls rapidly to the lower frequency level. The process which is responsible for this seems to be the excitation of the gas in the corona by disturbances caused by clouds of ionized material travelling outward at high speed.

Radio emissions can be detected from the Sun at all times, but they are subject to enormous ranges and variation. After continuing at a nominal level for many days there may be suddenly an increase in level of radiation many hundreds of times that of the normal level. Such radiations produce conditions called "noise storm". The largest of these usually follow the outburst of flares and the appearance of prominences. Many of these are associated with sunspots. Before the formation of a sunspot a magnetic field develops rather like a great magnet below the surface of the Sun; this may be as great as 30,000 miles in length. When this giant magnet reaches the surface of the Sun the magnetic field sticks up through the photosphere and activates the gas in the chromosphere. The gas is heated by the hydro-magnetic waves which travel up the magnetic lines of force. The spots then become apparent; they are dark because the temperatures at these points are lower than the surrounding regions and they mark the positions of the greatest concentration of the magnetic field. Above the spots the gas is hotter and there is violent activity; there are glowing clouds of gas which are suspended high up in the corona and these are supported by the magnetic field. During this time the flares may come into being.

This vast cloud of gas ejected at enormous speed drags the magnetic field out with it into the chromosphere; sometimes dragging the field out until it reaches the Earth. It gives rise to a wide range of radio emission as well as x-rays and cosmic rays.

The wave radiation can be divided into three parts: the first is the ultra-violet light which falls upon the ionosphere causing ionization of the "D" layer. The second gives rise to variations in the magnetic field of the Earth and causes fade-outs in radio communication between 5 megacycles and 20 megacycles, and the third associated with it is an interference on the very low frequencies of 10–15 kilocycles. Radio waves at 60 megacycles give rise to great outbursts of radio noise. All these, together with visible light, arrive at the Earth some eight minutes after the outburst has begun on the Sun. A later effect is that of cosmic rays created during the process, and, consisting as they do of atomic nuclei, again cause ionization in our upper atmosphere. This occurs some twenty to forty hours later than the appearance of the flare. At this time there are magnetic storms, ionospheric storms and, under very severe conditions, the appearance of the Aurora Borealis.

Wave radiations have simultaneous effects, but the delayed

effects caused by the particles may go on for some days after the origin of the outbursts. This whole sequence of events is illustrated in Fig. 8.

A particular study of the Sun was made by Christiansen and Mathewson in Australia. They built a large radio telescope consisting of a number of bowl-type aerials. With this instrument they scanned extremely small strips of the Sun. A comparison between the visible sunspots and the picture built up with this radio telescope showed that the radio waves came from areas very much greater than that of the visible sunspot.

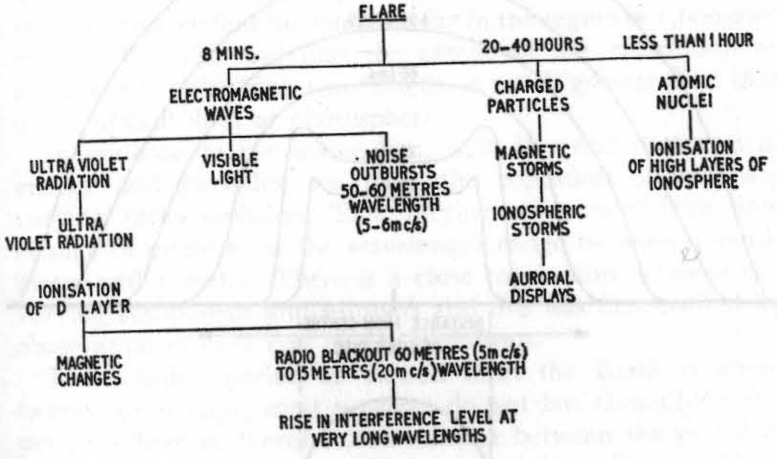


Fig. 8 Some of the effects of flares on the Sun

It showed conclusively how closely related the sunspots and radiations were. J. P. Wild and his associates built the first radio spectrograph for the study of these flares in Sydney in 1949 and by 1952 their apparatus had reached the stage where they were able to monitor continuously the radio spectrum between 25 megacycles and 210 megacycles. Using three rhombic aerials each covering a range of frequencies of two to one they were able, by means of a special receiver connected to these aerials, to follow the progress of a flare. Wild and his colleagues were able to recognize three types of bursts of activity. Type two and type three are similar although the type two bursts are by far the stronger. These are associated directly with the flare. Type one bursts do not closely follow after the activity caused by the flare. Near the sunspot maximum the flares increase in number and they may indeed occur at the rate of several in an hour. Some of these may last only a matter of seconds.

As opposed to the very active Sun at the time of these bursts,

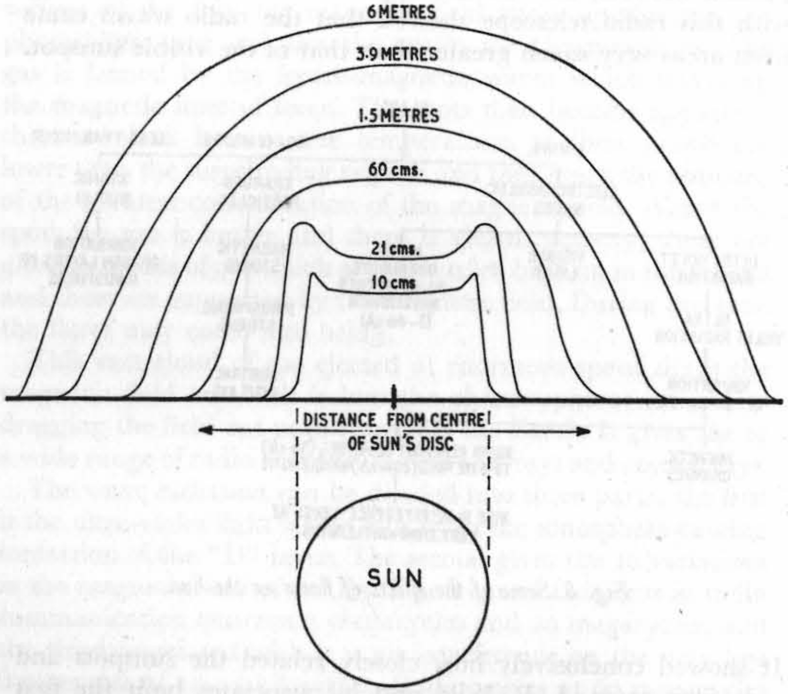


Fig. 9 Radio Sun. Extent of radio emission at various wavelengths

we have the condition of the quiet Sun. During this period most of the radiations are of thermal origin. The first observation of this solar radiation at wavelengths of a few centimetres was made with war-time radar equipment; this showed that there was a steady level of radiation which would correspond to a temperature of about 1,000,000° Kelvin. At centimetre wavelengths there is a pronounced increase in radiation in the regions at the edge of the visible disk. This is most

apparent between 10 and 30 centimetres wavelengths, thereafter it tends to be smoothed out with a more even distribution over the area of the visible disk. Fig. 9 shows the profile distribution of radiations at various wavelengths. This helps to give a clear picture of the shape of the radio Sun. The distribution of intensity across the radio Sun gives us a picture of the different levels of intensity of the corona and the chromosphere. Below 1 centimetre we see the chromosphere only and the temperatures are then of the order of 10,000° Kelvin. As we move out towards the corona, measurements at long wavelengths reveal the high temperatures that exist. At a wavelength of 3·7 metres we find the temperature in the region of 1,000,000° Kelvin. We also find that the extent of the radio Sun as measured at these long wavelengths is much greater than that of the optical disk, or photosphere.

In addition to the active Sun, with its great outbursts of energy and particles, we have the condition of a slowly varying radio emission. This varying component is a slow change of intensity in the wavelength range between 1 centimetre and 1 metre. There is a close connection between this varying component and sunspots and this was first noticed by observation of their day to day variations.

The rotation period as viewed from the Earth is about twenty-seven days; most sunspots do not last throughout this period. There is, therefore, no relation between the period of rotation and the variation of intensity of the radiations. However this is only part of the story, for it has been observed that the day to day variation extended over a period much greater than the life of the sunspots. It was possible to check this with observations made at the time of a solar eclipse. It was noted that as soon as the Moon moved between the Earth and the Sun the decrease of radio signals was not regular; the rate of variation suggested that there were a number of spots where radiation was more intense. Some of these areas coincided with sunspots but others were in no way connected with them. Waiting for solar eclipses is not a very satisfactory way to observe such phenomena.

Australian radio astronomers first investigated these areas of high level radiations by means of a radio telescope consisting of thirty-two paraboloids. This gave a very narrow beam some three minutes of arc in width and enabled a plot to be made across the disk of the Sun. The variation of the radiation showed that there were areas which would correspond to temperatures of 400,000° Kelvin to 1,000,000° Kelvin. Later measurements were made with a still larger radio telescope in the form of a cross-grating interferometer. This consisted of sixty-four paraboloids 19 ft in diameter arranged in the form of a cross; each arm was 1,200 ft long and this produced a series of narrow beams about three minutes of arc in width. The arrangement was such that they were separated by one degree between centres. The Sun, therefore, could be seen only by one beam at a time. With this instrument a very reliable picture of the distribution of radiation could be built up.

The total extent of the Sun's corona is not known. Though the Sun is wholly gaseous the Sun's atmosphere can only be said to end where the density has fallen to that of the gas of interstellar space. It would appear that in the area occupied by the Earth the density may be some 9,000 protons, that is, hydrogen nuclei, and 9,000 electrons per cubic inch. This is, of course, a very tenuous gas but still much denser than that of interstellar space; the actual temperature may be as high as 100,000° Kelvin. This high temperature is maintained by a still higher temperature of the corona itself; the gas near the Sun being hotter, particles moving away from it will have greater energy than those that are moving toward the Sun. Part of the energy of outgoing particles will be given up to those coming in and so heat will be passed across space by the process of conduction. This method of transfer of heat deals only with a small part of the heat; most of it is conveyed by means of electro-magnetic radiation both visible and invisible.

A certain part of this heat is conveyed by still another process, namely convection. Propulsion of energetic particles directly from the lower atmosphere of the Sun at the time of a violent disturbance, such as a flare, is the reason for this transfer. There is thus a direct contact between the atmosphere of the Earth and that of the Sun. Though conduction and convection are concerned with but a small fraction of this transfer of heat, nevertheless, the particles which are accelerated by solar storms can disturb radio communications to a very marked degree. Linking of the two atmospheres that gives rise to the

continued turbulence in the ionosphere are the radiation belts now well known as the Van Allen belts.

Each year an attempt is made to discover the extent of the radio Sun. During June each year the Sun passes close to the position of the Crab Nebula. The Crab Nebula is a source of very intense radio waves; since the Sun's corona should be able to absorb or modify the intensity of radio waves passing through it, it should be possible to detect the change in level of intensity from the Crab Nebula as it is eclipsed by the Sun. This measurement, first carried out by A. Hewish in 1951, confirmed the Sun did in fact obstruct these radio waves from the Crab Nebula. Fig. 10 shows the extent of the reduction of the radiations as the Sun passed across the area. Measurements are made some fifteen days before the actual time of eclipse and continue until a similar period after. Confirmation is not only received of the obstruction of the radio waves, but also that it is the lower frequencies which are affected first. This agrees with our other findings of the intensity of radiations from the corona, that is that the higher frequencies are nearer to the surface of the photosphere than the lower frequencies. It is therefore of considerable importance to our investigations to be able to conduct these observations of eclipses by the Sun of radio sources

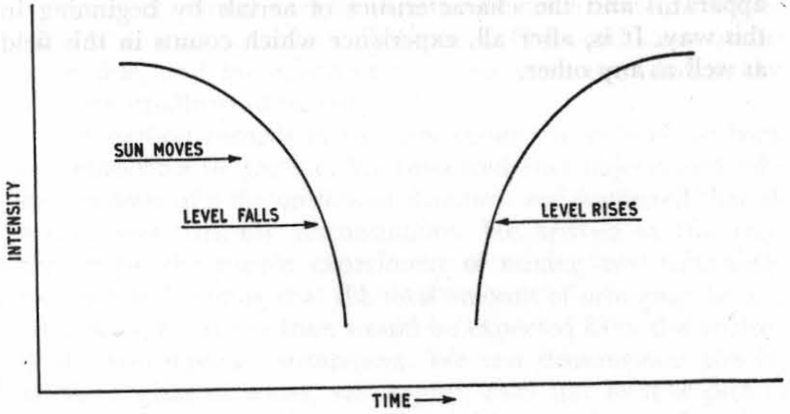


Fig. 10 Reduction in radiations from Crab Nebula as Sun's corona occults its position

throughout the period of the sunspot cycle. Though the Sun has been under continuous and intense investigation there is still much to be learned about the processes and their effects on outer space.

Other methods are now coming into operation for the study of the Sun's corona and efforts are being made to send powerful radar signals up to the corona in an attempt to get reflections from it. It is as yet too early to express an opinion as to the value of these techniques, but there is no doubt that sooner or later these experiments will prove successful.

The Sun offers the amateur an important subject with which to begin his attempts at radio astronomy. Comparatively simple apparatus is required, even the aerial can be very elementary. Much valuable experience can be gained about the use of apparatus and the characteristics of aerials by beginning in this way. It is, after all, experience which counts in this field as well as any other.

Chapter Five

BASIC ELECTRONICS–I

IT IS AN accepted scientific fact that all forms of matter are built up of molecules. Molecules are exceedingly small and can be described as the smallest portion of any piece of matter which can still retain the characteristic properties of the substance. We are unable to see these molecules even with the most powerful microscopes available. However, this does not prevent us from knowing quite a lot about them, and the science of chemistry enables us to break down these molecules into still smaller particles called atoms. The word atom means indivisible, and for many years atoms were regarded as the ultimate smallness obtainable.

Our earliest records of thinkers about this subject go back to Democritus in 400 B.C. He suspected that objects and substances were of a discontinuous structure and suggested that *all* matter was basically discontinuous. He arrived at this conclusion by the simple experiment of mixing two substances together and finding that the total amount of substance he had after mixing was less than would be expected from the volume of the two separate substances. We can demonstrate this by taking a glass of water, say a pint, and add to it a pint of alcohol. We shall find that the resulting mixture is less than 2 pints; this is because part of the alcohol has worked its way into the spaces between the water particles, and part of the water has worked its way between the alcohol particles. This happens in gases and in solids. Also if we take a bar of gold and put it in close contact with a bar of silver and press the two very firmly together, we shall find that after a period of time some of the gold has migrated inside the silver, and some of the silver has migrated to the inside of the gold. In a similar way solids can diffuse into other solids, gases fuse into other gases, and of course, solids dissolve in liquids.

We now know that atoms are still further divisible into constituent parts of electrons, protons and neutrons. If two or more similar atoms are combined together they form a molecule of one of the substances which we call elements. Originally we had a list of 92 elements, though of recent years this number has been increased.

The number of an atom in this series of elements is called its atomic number. Hydrogen is an element with an atomic number one; carbon is an element with an atomic number six; copper has an atomic number of twenty-nine; gold has an atomic number of seventy-nine, and bismuth an atomic number of eighty-three. The complete list of these elements can be found in any chemistry text book.

These elementary substances remain always the same. No matter how they are combined they never change their condition. They may be combined in many ways but they always retain their own identity and when a substance is analysed the original elements can be separated out. Substances which contain different atoms are called "compounds"; for example, water is a compound and has two hydrogen atoms and one oxygen atom and the formula is written H_2O. Sulphuric acid has two hydrogen atoms, one sulphur atom and four oxygen atoms and its formula is in consequence written H_2SO_4.

Until the latter part of last century our knowledge of atoms was theoretical. Many chemists felt sure that the concept of the atom as an ultimate particle was merely a useful fiction. It allowed a plausible explanation about the structure of matter but it obviously could not be true. The most important reason for doubting this conception of the atom was from the fact that atoms join with each other to form molecules. Atoms are not like billiard balls which, however many you put together, will not stick to each other, or in combinations of various numbers. Atoms do this and because they combine in this way there must be something special about their surfaces; there must be something which enables them to hold tightly on to one another. Chemists conveniently call this property "bonds". We can describe this bond as the capacity of an atom to hold the hydrogen atom or its equivalent. For example, an oxygen atom is able to hold two hydrogen atoms as in H_2O, which is water. The oxygen atom, therefore, has two bonds. A carbon atom can hold four hydrogen atoms and it, therefore, has four bonds. Now, since the oxygen atom has two bonds and the carbon atom has four, carbon can hold two oxygen atoms as for example, in carbon dioxide, which is CO_2. These bonds are spoken of as "valences". Hydrogen is singly valent, or monovalent, oxygen is di-valent, nitrogen tri-valent and carbon tetra-valent. Some bonds are tighter than others; if the bonds are loose the compounds are less stable than if the bonds are tight. It is, therefore, easier to take these less stable compounds and break them apart to release energy. Chemists were quite satisfied to know of these bonds without inquiring too closely as to what they actually were.

It is curious to note that the solution to this problem came from electricity and not chemistry. The Greeks knew that when an amber rod was rubbed with cloth or fur, the amber became electrically charged and could pick up bits of paper and pith. The Greeks called amber "elektron" and the name has survived to this day. This kind of electricity was called "static" or "frictional" electricity and for many years it was considered to be entirely different from electricity which flowed, such as that which had been discovered by Galvani and Volta in the eighteenth century. They used chemical batteries and caused electricity to flow along wires. In the nineteenth century a tube was invented called the "Geissler" tube. This was a glass tube with a metal plate sealed in at each end and connected by wires to an outside source of electricity. When the tube had been evacuated of air it was found that the electricity could jump across the gap between the two metal plates. Something must pass through the residual gas in the tube and flow from one electrode to the other. If a small piece of zinc-sulphide is placed in the path of the stream inside the tube the zinc-sulphide lights up the moment the current is turned on. It can be observed that the glow is made up of tiny little spots of light all over it. The bursts of light are all of the same size and this indicates that the electricity passing through the tube must be composed of small, uniformly-sized pieces. The stream that passes through the tube is made up of negatively charged units and their charge can be demonstrated by deflecting the stream by means of a magnet placed outside the tube.

J. J. Thomson, the English physicist, first threw light on this problem of electronic particles. He measured the magnetic force required to divert the stream of charged particles in the tube and by this means he was able to measure their mass. He discovered that they were very much lighter than anything previously known to science. He found that each of the charged units weighed only 1/1840th of a hydrogen atom. Thomson called these new particles "electrons" and demonstrated that they are, in fact, the electric current.

Electricity is composed of electrons, and when current flows it is the electrons which move. Thomson had therefore discovered a new fundamental particle, for electrons are all alike. They all weigh the same and they all have the same negative electrical charge.

The simplest atom in the atomic scale is that of hydrogen. It is the lightest element we know and consists of a central core containing one proton with one electron spinning in an orbit around it at high speed. It shows the exact balance of one electron with a negative charge balanced with one proton with a positive charge. This elementary positive charge, which we call the proton, is very heavy compared with the electron. The atom consists almost entirely of empty space. If we could make the atom as large as, say the Albert Hall, then the proton and electron would be scarcely larger than pins' heads in this vast expanse.

If we look at the helium atom, which is more complicated than the hydrogen atom, we find that its nucleus contains two protons and two neutrons. In order to balance this unit and make it electrically neutral it needs two electrons with a negative charge, and these two electrons are in orbits around the nucleus.

This knowledge enables us to understand how it is that we are able to create electricity by friction when we rub the amber with the piece of fur. The active friction detaches the orbital electrons from the atoms in the fur and transfers them to the amber. This means that the amber has now become negatively charged because it has collected more than its quota of negative electrons. The fur, on the other hand, has lost two electrons and has an excess of positive charge. We see, therefore, that the unbalance produced in these two materials provides positive and negative electricity. This picture is, of course, a very simplified one and in actual fact the atomic construction is much more complicated. For our purpose, however, it will be

sufficient to understand how the disorganization of these orbital electrons leads to the flow of an electric current.

A substance which has a number of free electrons wandering about in the inter-atomic spaces is a good conductor of electricity. Copper, for example, is an extremely good conductor of electricity. Raising the temperature of such a substance tends to increase the number of free electrons and if it is raised to white heat they may even escape altogether into surrounding space. This they do by virtue of the high speeds they attain as the substance gets hotter and hotter. Those electrons which are attached by strong bonds will still continue to keep their places.

A current of electricity can be caused to flow by applying a voltage across the substance. Most metals are conductors but they vary from one to another in the manner in which the current of electricity can flow. Some allow electrons easy passage, while in others flow of the free electrons is more difficult. This property in a substance is known as "resistance". A certain amount of energy has to be used in order to overcome this resistance, which is why the normal radiant heat fire which is connected to the supply mains consists of a conductor of high resistance. When the mains are applied to this high resistance the temperature is raised by the energy used up in moving the free electrons and it appears as heat which we use to warm the room.

When a substance is a very bad conductor it is called an "insulator" and among such substances are ebonite, glass, sulphur, mica, wax, plastic, and so on.

The conduction of electricity in metals is due to the flow of electrons. Good conductors are those containing many free electrons that are only lightly bound to the parent atoms and these are able to move from one atom to another under the influence of an electric field which has been set up by the application of a voltage, or potential, across the conductor. The direction of flow of electrons is opposite to the direction of flow in the conventional sense. It is usual to speak of current flowing from positive to negative, but this is, in fact, incorrect, for what really happens is that the positive end pulls the electrons while the negative end pushes them. Thus, through the conductor, electrons move from negative to positive and the source of the voltage or potential has its electrons moving from positive to negative.

We therefore have the complete cycle of the internal and external circuit. This is the simplest form of circuit and we shall be dealing with this in greater detail later on when discussing circuitry.

The conduction of electricity through solutions is rather different. In this case it is due to the movement of the ionized atoms, that is ions which have gained or lost electrons. Both positive and negative carriers are present in this case and they move in opposite directions when a field is created by an applied voltage.

Another important aspect of the conduction of electricity, so far as radio astronomy is concerned, is the conduction of electricity in gases. This conduction can be due to positive and negative carriers. Some of the negative carriers may be free electrons which are not attached to the atom. Free electrons can be accelerated in conditions of high magnetic field and if this happens in a gas they may collide with another atom of gas with such energy that they tear off the outer electrons forming other conducting particles. A strong current can be built up by this process of multiplication, some of the electrons may be recaptured by electron-deficient ions and energy is, in this case, liberated in the form of radiation. Should this occur where there is sodium present the radiation emitted would be that of visible yellow light.

In electronic circuits we are concerned with the vacuum tube, or valve. Electrons are used in this case after they have been evaporated from the surface of the heater which is energized inside the valve. A potential applied to the other electrodes in the valve causes a current to flow, the magnitude of which is determined by the external circuits. These properties are the bases of our receivers and other electronic equipment used in radio astronomy.

In recent years, semi-conductors in the form of transistors and diodes have come into universal use. The current in these cases is rather more complicated in the manner of its operation. It is partly carried by relatively free electrons and partly by positive carriers. In the case of semi-conductors a positive carrier is really a gap in the lattice, or pattern, of the tightly meshed electrons in the crystal structure. This gap is caused by the absence of one electron. Now, if an electron from a neighbouring atom on one side of the gap moves along under the influence of an electric field it will fill the gap which has been left and in consequence the gap will have appeared to move one space. Successive movements of electrons will result in the apparent gap moving continuously in one direction. Since this is in the opposite direction to that of the negative electron it behaves like a positive charge. We shall deal with this in more detail later.

We have spoken of a battery as a source of electric potential and we use the term "potential difference" and "positive" and "negative". The battery is one form of apparatus with which we develop electro-motive force, which is abbreviated "EMF". This term EMF refers to the pressure generated in the interior of the source. This results in the production of a potential difference between the terminals. A battery is one form of this source, a generator or dynamo is another.

We commonly speak of two kinds of current, direct current and alternating current.

Direct Current

A cell, or battery of cells, is a chemical means of producing electro-motive force. This simplest form of cell can consist of a container of glass of dilute sulphuric acid having immersed in it two electrodes, one of copper and the other of zinc. If wires are attached to these two electrodes and the ends coupled to a small torch bulb a potential difference between the two electrodes will cause a current to flow through the lamp which will then light up.

In practice this very simple cell has certain disadvantages, and more highly developed and sophisticated types of cells are in use, but the principle is the same in all cases. These cells produce direct current. Three kinds of battery are now current, the dry battery and two forms of wet battery. The most familiar form of wet battery is that of the lead-acid accumulator which is used on the ordinary motor car. This has the advantage that it can be continually recharged so that many successive cycles of discharge and recharge can be maintained throughout its useful life. A dry battery, after a normal specified period of time and use, becomes chemically useless and has to be scrapped. The other type of wet battery also has a limited life but this is not met with very much nowadays. It is called a Leclanché cell. It consists of a plate of carbon and zinc in a solution of salammoniac. Each cell itself has a comparatively low voltage and a number may be combined in series to achieve a higher potential difference between the positive and negative terminals. Groups of batteries in series may also be combined with other groups in parallel to allow larger currents to flow.

We have spoken of voltage, current and also resistance. Each of these terms bears a definite relation one to the other and this relationship is called "Ohm's Law". This law in its simplest form for D.C., or direct current, is "The current in a conductor is directly proportional to the applied voltage"; this we may write in the following form:

$$\frac{\text{voltage V}}{\text{current I}} = \text{resistance R.}$$

Similarly, we can transpose this and say

$$I = \frac{V}{R} \quad \text{and} \quad V = I \times R.$$

This enables us to work out one of the quantities if we know the value of the other two. This law is also true for all metallic conductors.

Circuit Theory

Electronic circuits consist principally of resistance, inductance and capacity. Resistors in a circuit may be connected either in series or in parallel or a combination of both. When they are in series they can be added up in a normal manner so the total resistance will be the sum of the number of resistances in series, that is $R = R_1 + R_2 + R_3$ and so on. The current flowing through these resistors therefore would be, applying our formula:

$$I = \frac{V}{R} = \frac{V}{R_1 + R_2 + R_3}.$$

When conductors are connected in parallel then they are added in a rather different manner, that is

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \text{ etc.}$$

Here the total resistance will be less than the lowest resistance in the circuit. If the circuit consists of some parallel and some series elements then the parallel group will be added to the series group, see Fig. 11.

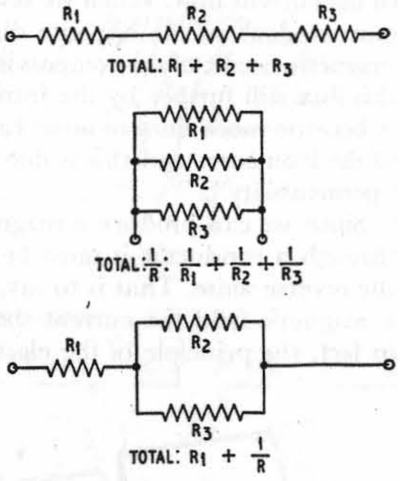


Fig. 11 Series and parallel resistors

When a current passes along a conductor a magnetic field is automatically created round that conductor. The direction of the lines of magnetic force, as they are called, will be related to the direction in which the current is flowing. Near the surface of the wire the density of this magnetic flux is greatest, and as we move further and further away at right angles the flux gets less. If we wind the conductor in the form of a solenoid, Fig. 12, the magnetic field due to these turns increases the flux considerably. We can test with a compass needle to indicate the direction

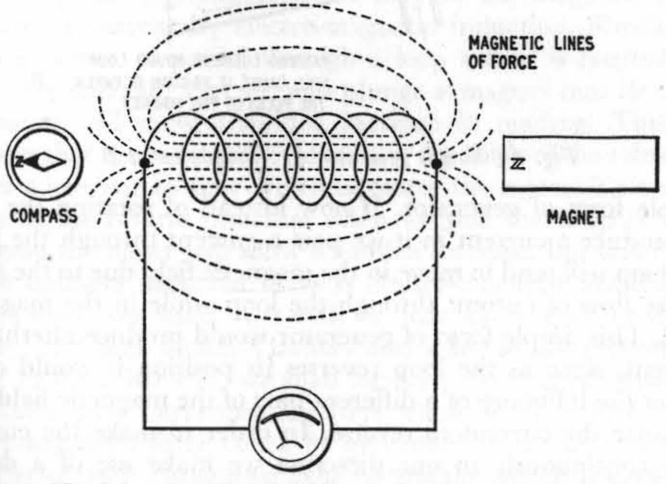


Fig. 12 Increase of magnetic field by winding turns of wire in the form of a solenoid

of the current flow. When we reverse the connection so that the current flows in the opposite direction, we shall find that the magnetic needle of the compass is also reversed. We can increase this flux still further by the introduction of an iron core. This is because more lines of force can be concentrated in the area of the iron core, and this is due to a quality of the iron called "permeability".

Since we can produce a magnetic field by passing a current through a conductor it must be possible to make this work in the reverse sense. That is to say, if a conductor passes through a magnetic field the current should be induced in it. This is, in fact, the principle of the electric generator. Fig. 13 shows a simple form of generator. If now instead of rotating the loop to produce a current in it we pass a current through the loop, the loop will tend to move in the magnetic field due to the effect of the flow of current through the loop while in the magnetic field. This simple form of generator would produce alternating current, since as the loop reverses its position it would come under the influence of a different part of the magnetic field and so cause the current to reverse. In order to make the current flow continuously in one direction we make use of a device known as the "commutator". This is illustrated in Fig. 14.

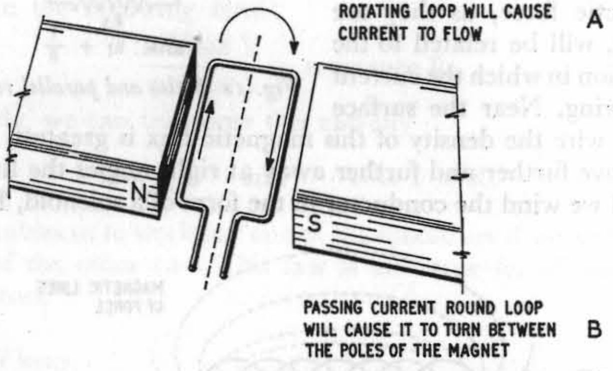


Fig. 13 Simple generator A. Simple motor B

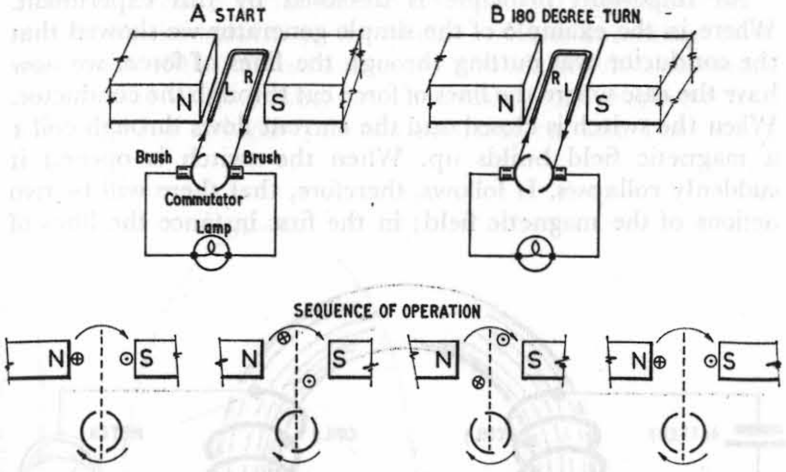


Fig. 14 Commutator to produce one way (D.C.) current

This simple type of generator is the basis of many of our instruments which can be considered as small motors working within a limited range. If we attach a pointer to the loop the strength of the current passing through it will cause the pointer to move a certain distance and this we can calibrate so as to indicate the value of the current flowing. In the case of the single loop, the movement of the loop in the magnetic field produces a current by electro-magnetic induction. Similarly, when a current is passed through a loop a field is created. If we take a solenoid, Fig. 12, and plunge a magnet into its coils the meter will show a sudden increase in reading. This illustrates that the magnetic field cutting the turns of the solenoid causes a current to flow which actuates the meter. This also works in the reverse direction when we pull the magnet out, for again the meter will show a sudden increase, but this time in the opposite direction. This is electro-magnetic induction.

Supposing we wind two coils of wire on a core of iron, Fig. 15, and then attach a battery and a key to one coil and a meter to the other coil, we shall have an example of Michael Faraday's early experiment in 1831 when he discovered that electro-magnetic force could be generated by an inter-action of a circuit and a magnetic field. When the switch is closed a current flows through the coil at 1, this will immediately cause the meter to move; when the switch is released and the current broken the meter will again show movement, but this time in the opposite direction.

An important principle is disclosed by this experiment. Where in the example of the simple generator we showed that the conductor was cutting through the lines of force, we now have the case where the lines of force cut through the conductor. When the switch is closed and the current flows through coil 1 a magnetic field builds up. When the switch is opened it suddenly collapses. It follows, therefore, that there will be two actions of the magnetic field; in the first instance the lines of force will cut the conductors as the current flows. In the second case, as the circuit is opened the lines will collapse and in collapsing cut the conductors again. This accounts for the meter moving first in one direction and then in the other. The important thing to remember is that the field has grown and collapsed.

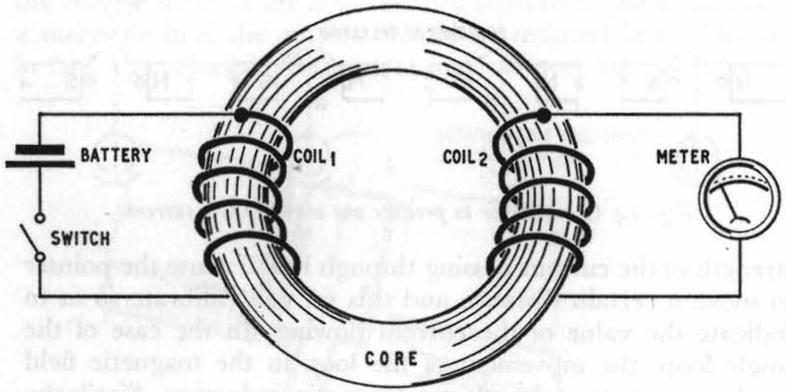


Fig. 15 Two coils of wire wound on a magnetic core to demonstrate electro-magnetic induction

We can do this by another method, and that is to use alternating current. In direct current the electrons flow in one direction only under the influence of the potential difference. With alternating current the electrons are caused to flow first in one direction and then in another. This flowing backward and forward can be slow or fast, and the number of times that it reverses in a second is called the "frequency" or "periodicity". The normal household supply mains have a periodicity of 50 cycles a second. In electronics we deal with frequencies up to many millions of cycles per second. The principle, however, is the same; the current in the circuit reverses according to the frequency of the alternating current. This provides us with a simple way of increasing and decreasing the magnetic field. Fig. 16 shows a single cycle of alternating current. The current builds up toward the left to a maximum, decreases to the zero point, then increases on the right to a maximum, and then again decreases to zero. Thus the magnetic field will first be caused to build up, then collapse, then build up in the opposite direction and then collapse again.

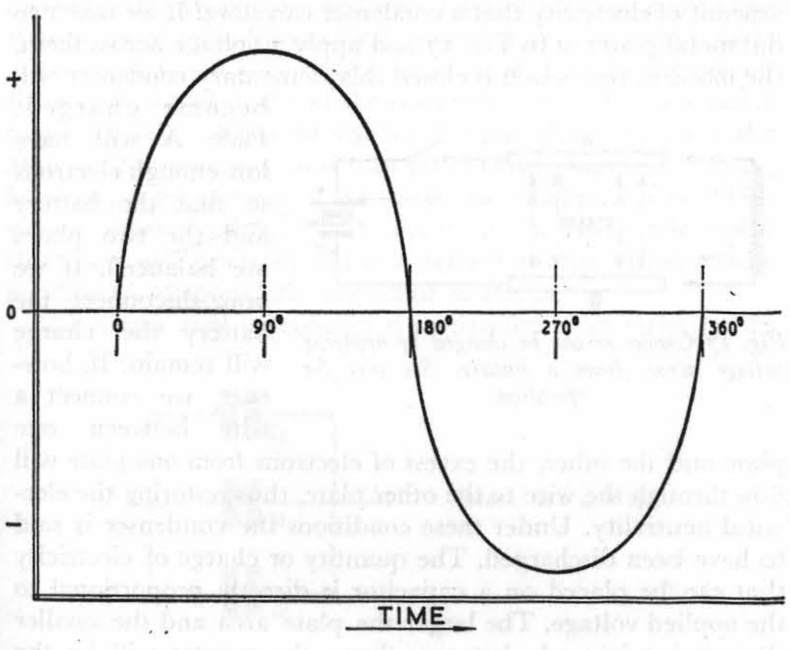


Fig. 16 Single cycle of alternating current

If we now go back to our two coils of wire in Fig. 15, and we replace the battery and the switch by an alternating current generator, we shall find that when the alternating current flows in coil 1 the meter in coil 2 will move backward and forward in sympathy with the variation of the current in coil 1. We are thus able to transfer energy from one coil to the other without any direct connection between them. A device of this kind is called a transformer. The importance of this principle will appear later when we come to discuss circuits which have coils in them for special purposes.

We have dealt with resistance and inductance and we must now deal with capacity. Capacity is the term applied to the amount of electricity that a condenser can store. If we take two flat metal plates as in Fig. 17 and apply a voltage across them, the moment the switch is closed this elementary condenser will become charged. Plate A will have lost enough electrons so that the battery and the two plates are balanced. If we now disconnect the battery the charge will remain. If, however, we connect a wire between one plate and the other, the excess of electrons from one plate will flow through the wire to the other plate, thus restoring the electrical neutrality. Under these conditions the condenser is said to have been discharged. The quantity or charge of electricity that can be placed on a capacitor is directly proportional to the applied voltage. The larger the plate area and the smaller the spacing is made between them, the greater will be the capacitance of the condenser.

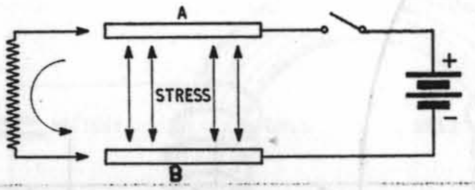


Fig. 17 Condenser can be charged by applying voltage across from a battery. See text for operation

The capacitance also depends upon the medium between the plates. Certain insulating materials have a property of increasing the capacitance many times. The measure of this ability is called the "di-electric constant", for when an insulator is placed between the two plates of the condenser this is called a di-electric. When the di-electric is air the capacitance is at its lowest. Capacitors used in electronic apparatus vary in size, construction and their capacity. Some, such as tuning condensers, may be air-spaced and variable, others are of fixed values in the circuit. Fixed capacitors are assemblies made of thin plates of metal foil with a di-electric between them. Extremely large capacitances can be provided in quite small units. Solid di-electrics are most commonly mica, paper or special ceramics.

Another type of capacitor is the electrolyte capacitor. This uses aluminium foil plates with a conducting-liquid chemical between them. The di-electric is a film of insulating material that forms on one set of plates due to electro-chemical action when a D.C. voltage is applied to the capacitor. Very high capacitances can be obtained with a given plate area with this type of condenser. When a high voltage is applied to the plates of the capacitor a very considerable force becomes exerted on the electrons and nuclei of the di-electric. Since the di-electric is an insulator, electrons cannot become detached from the atom in the same way that they do in conductors. If the force, however, is great enough the di-electric can break down. This usually takes the form of a puncture in the di-electric which thereafter will probably allow a current to flow. Under these conditions, of course, the capacitor is useless.

Capacitors may be connected in parallel or in series just in

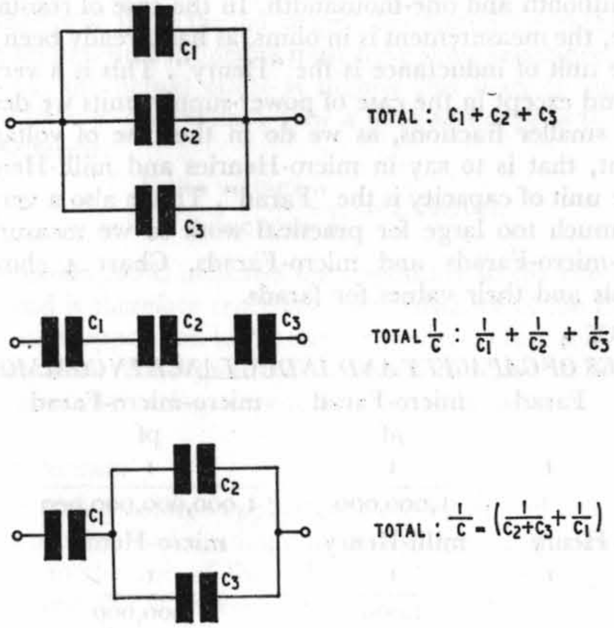


Fig. 18 Condensers in parallel, series and series parallel

the same way as resistors. However, the method of adding the capacitors is slightly different. Capacitors in series add up to a total which is less than the smallest capacitor, and in this respect it is the same as the parallel resistor.

In the case of condensers in parallel they are added together in the same way as series resistors are added. The total capacity is the sum of all the capacities, and this is illustrated in Fig. 18.

Units

So far although we have spoken of voltage, current, resistance, inductance and capacity, we have given no indication of the value of the units used with these components. Voltage, current and resistance are closely related, as we have seen from the formula of Ohm's Law. Voltage, as its name implies, is measured in volts, current is measured in amperes. When dealing with electronic circuits small fractions of a volt and an ampere are in common use and we deal with micro-volts and milli-volts, micro-amps and milli-amps. These are respectively one-millionth and one-thousandth. In the case of resistance, of course, the measurement is in ohms, as has already been stated.

The unit of inductance is the "Henry". This is a very large unit and except in the case of power-supply units we deal with much smaller fractions, as we do in the case of voltage and current, that is to say in micro-Henries and milli-Henries.

The unit of capacity is the "Farad". This is also a very large unit, much too large for practical work so we measure it in micro-micro-Farads and micro-Farads. Chart 4 shows the symbols and their values for farads.

CHART 4

VALUES OF CAPACITY AND INDUCTANCE IN COMMON USE

Farad	micro-Farad μf	micro-micro-Farad pf
1	$\frac{1}{1,000,000}$	$\frac{1}{1,000,000,000,000}$
Henry	milli-Henry	micro-Henry
1	$\frac{1}{1,000}$	$\frac{1}{1,000,000}$

When speaking of inductance we mentioned alternating current. We must deal with it now in greater detail. When an alternating voltage is applied to a resistance a current flows exactly in step with the voltage, that is to say the voltage and the current are in phase. This is true at any frequency of the current, providing that it is a pure resistance. At very high frequencies, such as radio frequencies, even a resistor may have certain other effects which are called "reactance". If the circuit is purely resistive then Ohm's Law will apply just the same for alternating current as for direct current.

Reactance

Reactance is of two kinds in alternating current circuits. Capacitive reactance and inductive reactance. When an alternating voltage is applied to a pure inductance, that is to say, one with no resistance, the current is 90 degrees out of phase with the applied voltage, in other words, the current lags 90 degrees behind the voltage. The reason for this is that the back EMF generated by the varying voltage opposes the rise of current in the inductance causing it to lag behind. If the inductance was pure the 90 degree lag would be exact, but because all practical inductors must have some resistance, however small, the actual effect is slightly less than 90 degrees. Here we have to vary Ohm's Law a little and express it in the following terms:

$$\frac{\text{E the voltage}}{2\pi\text{F} \times \text{inductance L}} = \text{I current.}$$

The expression $2\pi FL$ indicates the opposition to the rise of the current and it therefore represents the resistance of the circuit. The degree of opposition to the rise of the current is proportional to the frequency; the higher the frequency the more difficult it is to drive an alternating current through an inductance.

Capacitive Reactance

An alternating voltage applied to a condenser will result in a current which leads the voltage by 90 degrees. In this case the opposition to the rise of the voltage is due to the capacitive reactance and this is expressed as:

$$\frac{1}{2\pi FC}.$$

The formula for this type of circuit then becomes

$$\text{I current} = \frac{\text{E voltage}}{\frac{1}{2\pi FC}}.$$

Again the frequency is important and the higher the frequency the easier does A.C. current flow in a condenser.

We must now consider circuits which contain all three of these conditions, resistance, inductance and capacity. This results in a somewhat complicated formula which is expressed as:

$$\text{I current} = \frac{\text{E voltage}}{\sqrt{\text{R}^2 \text{ the resistance} + \left(2\pi FL - \frac{1}{2\pi FC^2}\right)}}.$$

The value under the square root sign represents the total opposition that is in the circuit, that is to say the resistance and the reactive parts, and when dealing with alternating currents this is called the "impedance" of the circuit. We measure impedance in ohms and usually designate it by Z. We can now rewrite our formula for A.C. circuits and say:

$$\text{I current} = \frac{\text{E voltage}}{\text{Z impedance}}.$$

With an alternating current circuit containing resistance and capacitance a very important effect takes place when the inductive-reactance and the capacitive-reactance just balance one another. This means that the impedance of the circuit is reduced to that of the resistance alone since the reactances cancel one another. This happens at a particular frequency and is called the "resonant" frequency. This is a most important part of the property of a circuit, for it enables us to tune our circuits to respond to a particular frequency in the electromagnetic spectrum. In the next chapter we will deal with circuits making use of the properties we have discussed in this chapter.

Chapter Six

BASIC ELECTRONICS-II

IT WAS STATED in the previous chapter that in electronic circuits we are concerned with the vacuum tube or valve. The valve is a device in which we use the effects of thermionic emission. If a wire is heated to incandescence in a vacuum the electrons near the surface are stimulated into motion and fly off into the surrounding space. The higher the temperature the greater is the number of electrons that will be emitted. We call this source of electrons a "cathode". This cathode is specially coated with material which is rich in free electrons. If the cathode is the only electrode in the valve, most of the electrons thrown off will stay in the vicinity of the cathode and there they will form a sort of cloud. There is a reason for this. Electrons which have been thrown off are negative particles of electricity and they form a negative charge which we call a "space charge". This space charge prevents those electrons nearest the cathode from getting too far out because, as we have previously seen, like charges repel one another, so that the electrons in the space charge tend to push the free electrons back on to the cathode.

Now, if we introduce a second electrode, as we did in the case of the Geissler tube, and apply a potential between the cathode and the other electrode, which we call the "anode", and make the anode positive with respect to the cathode, then a current of electrons will flow from the cathode to the anode. If, therefore, we connect a battery to a two-element thermionic tube, or valve, to which we give the name of "diode", a current can be made to flow in the circuit, Fig. 19.

Since the electrons are particles of negative electricity they will be attracted only to the electrode when it is positive with respect to the cathode. If we should give this anode a negative charge all the electrons will be repelled back to the cathode and no current will flow. From this it will be clear that the diode is a one-way device, that is to say, we can make electrons flow from the cathode to the anode, but we cannot make them flow from the anode to the cathode. This is a very important property for it enables us to use it as a gate or valve. We are thus able to use the diode in a supply circuit to change alternating current into direct current. It will be remembered that alternating current went first to a positive maximum, then to zero, then to a negative maximum and back to zero again. Consider then what would happen if we applied an alternating current across the diode instead of the battery. The anode would first rise to positive, in which case electrons would flow; it would then decrease to zero and the flow would stop and remain cut off during the next half-cycle when the voltage was negative. When again the voltage became positive, current would flow.

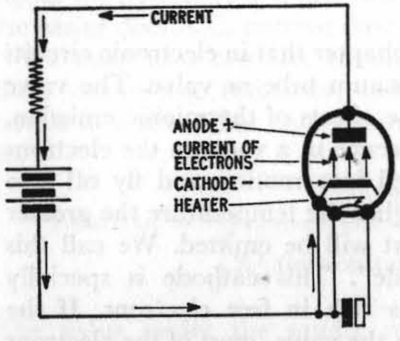


Fig. 19 Current in diode circuit

It will, therefore, be clear that current flows for only half the cycle, that is the positive half. The current flowing in this manner is said to be rectified alternating current. What we have in fact produced is an intermittent direct current, or one-way current. If there is now introduced into this circuit a resistance this will have pulses of direct current passing through it. The pulses can be reduced if a condenser is connected across the resistor, for then the condenser will charge up and in consequence tend to smooth out the variations caused by the intermittent pulses of direct current. This process is called "half-wave rectification".

It is possible to use this property of rectification in a number of ways. It can be used to provide direct current power supplies to operate our electronic apparatus by feeding a rectifier circuit with alternating current. It can also be used to rectify alternating currents which are produced in a radio receiver by a signal arriving at the aerial. We can in fact use the diode wherever we want a one-way passage of current.

The later development of the diode is known as a "triode".

It was discovered that if a third element was introduced between the cathode and the anode, and a negative voltage applied to this, the number of electrons flowing from the cathode to the anode could be controlled. In Fig. 20 is shown a simple triode valve. There is a battery for heating the cathode, a battery to supply a positive potential to the anode and a battery to provide a negative voltage to the third electrode, or "grid", as it is called. The grid having a negative voltage on it controls the space charge which surrounds the cathode. With a positive voltage at the anode there will be a point at which the voltage at the grid just stops the flow of electrons. If now this negative voltage is reduced some of the electrons will flow, and the more it is reduced the greater is the number of electrons allowed to flow from the cathode to the anode.

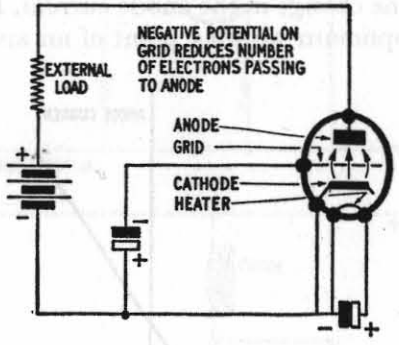


Fig. 20 Triode Valve

We now have a device where by varying the voltage on the grid we are able to control the number of electrons which flow from the cathode to the anode. This is a most important property and is known as "amplification". By a small change in voltage on the grid we can cause a large change of current inside the valve which will flow through the external circuit. The ratio of these two quantities is called "the amplification factor" and it is commonly designated by the Greek letter μ.

The amplification of a triode varies between three and more than a hundred. If, for example, the amplification factor is 30, then this would mean that if the grid voltage was changed by one volt the effect on the current flowing through the valve and the external circuit would be the same as if the positive voltage at the anode had been changed by 30 volts.

There is another definition which we use to indicate the effectiveness of a valve as an amplifier. It is a property called "mutual conductance". This takes into account the amplification factor and the anode resistance. The anode resistance varies with the amplification factor and where the amplification factor is low the anode resistance of the valve is relatively low. When it is high the anode resistance is relatively high. Thus mutual conductance can be described as a figure of merit for the valve. We arrive at its value by dividing the change in the anode current by the change in the grid voltage that causes the change in the anode current. Fig. 21 is a graph showing the optimum working point of an amplifier.

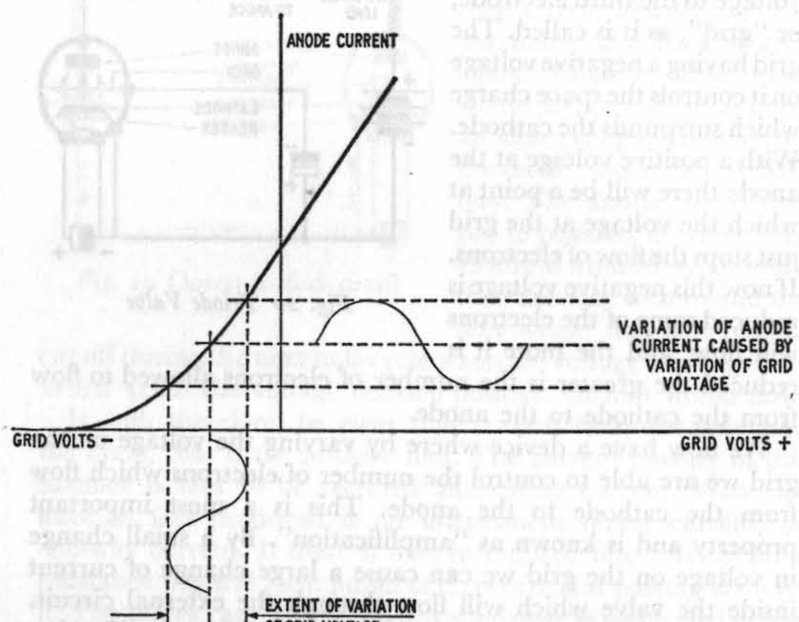
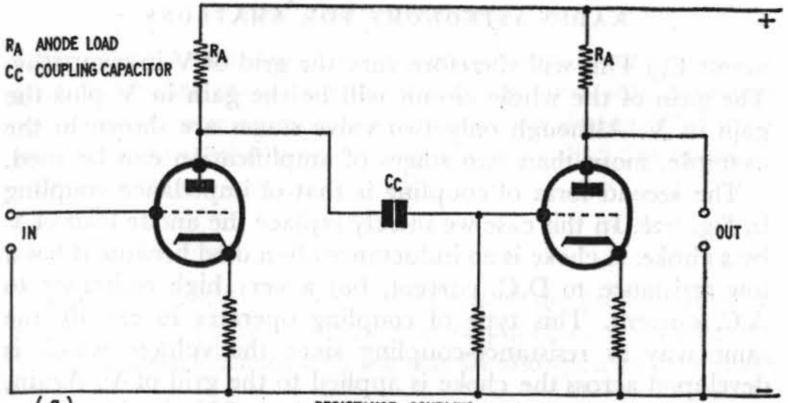
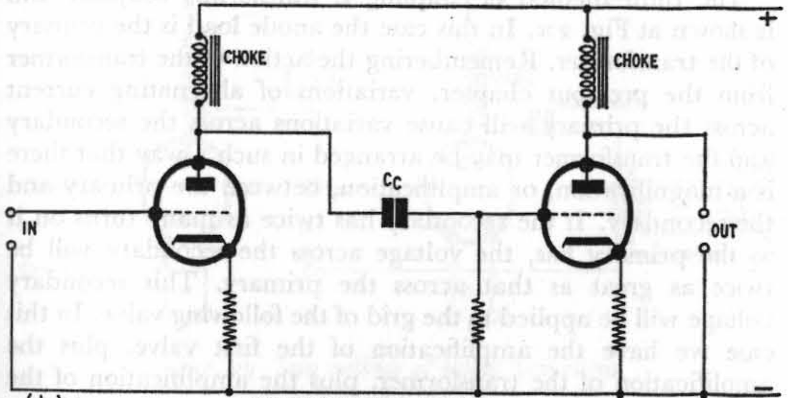


Fig. 21 Graph showing optimum working point of an amplifier

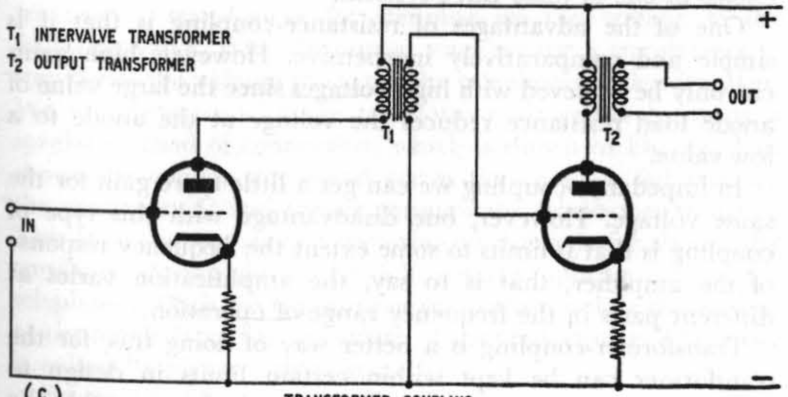
More than one amplifier may be coupled together and there are various methods of doing this. One form is known as "resistance coupling", another is "impedance coupling" and a third "transformer coupling". Simple circuits of this kind using two valves are shown in Figs. 22*a*, *b* and *c*. Fig. 22*a* shows a resistance-coupled amplifier. Here the voltage developed across the anode resistance R_A of V is applied via the condenser C_C across the following resistor R_A, coupled to the grid of the valve V_2. The alternating voltage set up across R_A will appear



(a) RESISTANCE COUPLING



(b) IMPEDENCE OR CHOKE COUPLING



(c) TRANSFORMER COUPLING

Fig. 22 (a) (b) (c) Three simple 2-valve circuits with different forms of coupling across C_C. This will therefore vary the grid of V in sympathy. The gain of the whole circuit will be the gain in V plus the gain in V. Although only two valve stages are shown in the example, more than two stages of amplification can be used.

The second form of coupling is that of impedance coupling in Fig. 22*b*. In this case we merely replace the anode load of V by a choke. A choke is an inductance often used because it has a low resistance to D.C. current, but a very high resistance to A.C. current. This type of coupling operates in exactly the same way as resistance-coupling since the voltage which is developed across the choke is applied to the grid of V. Again, the amplification of the two stages is that of V plus V.

The third method of coupling is transformer-coupling and is shown at Fig. 22*c*. In this case the anode load is the primary of the transformer. Remembering the action of the transformer from the previous chapter, variations of alternating current across the primary will cause variations across the secondary and the transformer may be arranged in such a way that there is a magnification, or amplification, between the primary and the secondary. If the secondary has twice as many turns on it as the primary has, the voltage across the secondary will be twice as great as that across the primary. This secondary voltage will be applied to the grid of the following valve. In this case we have the amplification of the first valve, plus the amplification of the transformer, plus the amplification of the second valve. There may be a number of reasons for making a choice of one of these three systems.

One of the advantages of resistance-coupling is that it is simple and comparatively inexpensive. However, high gains can only be achieved with high voltages since the large value of anode load resistance reduces the voltage at the anode to a low value.

In impedance-coupling we can get a little more gain for the same voltage. However, one disadvantage with this type of coupling is that it limits to some extent the frequency response of the amplifier, that is to say, the amplification varies at different parts in the frequency range of operation.

Transformer-coupling is a better way of doing this, for the transformer can be kept within certain limits in design to ensure that the transfer of various frequencies is reasonably the

same. Transformer-coupling is used when comparatively low-gain valves are used. If transformers are used with high-gain valves there arise difficulties in transformer design. The transformers used in this type of coupling are called "inter-valve" transformers.

Power Output Stage

In order to make audible the effects of the amplification we need an output stage to operate a telephone or a loudspeaker. In Fig. 23 two methods of doing this are shown: *a* for telephones and *b* for telephone or loudspeaker. In the case of amplification we used a transformer for stepping up the voltage to the following valve. With the output stage we use a transformer to step down the voltage to match the telephones or loudspeaker. We can dispense with the transformer altogether by using a special method of connection, which is shown in Fig. 24. This shows the voltage developed across the choke is fed through the condenser C and thence through the telephones to earth. This is not a very satisfactory way of doing it for should the condenser break down the full voltage appears across the telephones. With the transformer method we isolate the voltage at the anode from the secondary of the transformer. This stage is called "a power output stage".

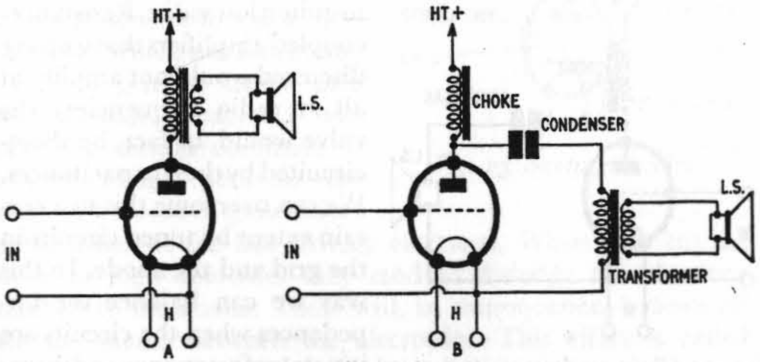


Fig. 23 Two methods of output connections

Power amplifiers are usually of low gain but are arranged to have large current changes. So far as radio astronomy is concerned only moderate output powers will be required in the monitor sections of receivers.

Thus far we have dealt only with audio frequencies, that is with the triode valve working at low frequencies. It was found that the ordinary triode valve had certain disadvantages when used in circuits, particularly radio frequency circuits which are circuits operating at the high radio frequencies and used to convey signals from the transmitter to the receiver. The electrodes in a valve act as small condensers. At radio frequencies the reactance of even the small inter-electrode capacities fall to quite a low value. Resistance-coupled amplifiers that we have discussed would not amplify at all at radio frequencies; the valve would, in fact, be short-circuited by these capacitances. We can overcome this to a certain extent by tuned circuits in the grid and the anode. In this way we can balance the impedances when the circuits are in a state of resonance and thereby achieve transfer of energy from one circuit to another. There are, however, limits to the amplification that we can obtain with triode valves even when we treat them specially with what are called "neutralizing circuits".

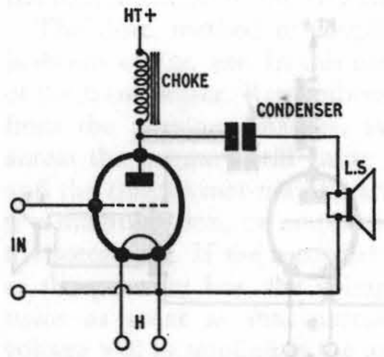


Fig. 24 Choke condenser output stage

At certain frequencies triode valves do have an advantage, and this we shall deal with later.

We must describe now the other various types of valves which are used when dealing with radio frequencies. The first of these is the screen-grid valve. We can reduce the capacitance between the grid and the anode by inserting a second grid between the control grid and the anode, as in Fig. 25. This second grid we call the "screen grid". It acts as an electro-static shield to prevent coupling between the control grid and the anode. This screen grid is usually in the form of a coarse mesh so that electrons can pass freely through it. There is, however, a

shielding action of the screen and unless it is positively charged the anode cannot attract electrons from the cathode as it does in the case of the triode. If a positive voltage is applied to the screen it attracts the electrons in the same way that the anode does in the triode. Most of the electrons attracted to the anode have sufficient velocity to get through the screen, but a certain number of them hit the screen itself with the result that some current will flow through the screen grid circuit. In order to isolate this grid from the rest of the circuit there is usually a capacitor to convey any unwanted radio frequencies to earth.

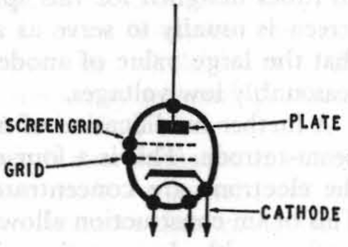


Fig. 25 Screen grid valve

A valve which has a cathode, a control grid, a screen grid and an anode is called a "tetrode". A certain condition is brought about between the anode and screen of these valves by the effect of the fast moving electrons. When they hit the anode at high velocities they tend to dislodge the electrons attached to the anode. These will, in consequence, bounce off into the spaces between the electrodes. This effect is called "secondary emission". In the triode valve the negative element, the grid, repels these secondary electrons back to the anode and they give no trouble, but in a screen-grid valve the positively charged screen attracts these secondary electrons causing a reverse current to flow between the anode and the screen. In order to overcome the effects of secondary emission it is usual to introduce another grid which is called the "suppressor grid". This extra grid inserted between the screen and the anode acts as a shield so that the secondary electrons cannot be attracted by the screen. They are therefore diverted back to the anode without seriously obstructing the normal anode current flow. This, then, is a five-element valve and is called a "pentode".

Although the screen grid in these valves partly reduces the influence of the anode on the total current flow, the control grid is still sufficiently able to control the anode current in essentially the same way that it does in the simple triode.

The additional arrangements of electrodes, however, does change the characteristics of the valve and they tend to increase the amplification factor and the anode resistance to a very high value. Because of this very high anode resistance the actual voltage amplification which is obtainable is much less than the amplification factor would seem to suggest. Although originally developed for application as radio frequency amplifiers, pentodes and tetrodes are used in audio frequency power amplifiers. In tubes designed for this special purpose the function of the screen is usually to serve as an accelerator of the electrons so that the large value of anode current can be made to flow at reasonably low voltages.

A further modification of multi-element valves is that of the beam-tetrode. This is a four-element valve so constructed that the electrons are concentrated in beams toward the anode. This beam construction allows very large currents to be drawn at reasonably low anode voltages. For power amplification, tetrodes are favourite valves and have largely supplanted other types in modern amplifiers.

There is another quality of multi-element valves which we must mention for it concerns the radio frequencies particularly. It was found that by varying the spacing of the mesh of the grid different values of grid bias could cause a variable change of characteristic. To this type of valve the term "variable μ" is applied. A variable μ valve can handle much larger signals than the normal type of valve, which usually has a sharp cut-off. It operates, therefore, as a variable amplifier which can be regulated according to the strength of the signal.

There are two other types of amplifiers with which we shall deal, that of the grounded grid and the cathode follower. These will both be described later when discussing special circuits.

Feed-back

When a part of the amplified signal is taken from the anode circuit of an amplifier and is returned to the grid circuit, feed-back is introduced. When the voltage inserted into the grid circuit is 180 degrees out of phase with the signal the feed-back is called negative, or degenerative. If the voltage is fed back in phase then it becomes positive, or regenerative. When the

feed-back is negative the voltage that is fed back opposes the signal voltage. The effect of this is to decrease the amplitude of the voltage which acts between the grid and the cathode and has the effect of reducing the overall voltage amplification. The amplifier, therefore, requires a larger exciting voltage in order to obtain the same output voltage and flow of current in the anode circuit. The greater the amount of negative feed-back applied to the circuit, the more independent the amplification becomes of either the circuit conditions or the valve characteristics. By this means a more constant degree of amplification can be obtained over the whole band of frequencies that is being used. The frequency response characteristic of the amplifier is then said to be flat.

Another advantage of feed-back is that distortion generated in the anode circuit of the valve tends to cancel itself out. Negative feed-back amplifiers are therefore reasonably free of harmonic distortion. Positive feed-back increases the amplification because the feed-back voltage is added to the original voltage and in consequence the larger excursions of the grid voltage cause a larger output voltage in the anode circuit. Usually the amplification will be greatest at one particular frequency. This will depend upon the circuit arrangements. If enough energy is fed-back the circuit will go into a state of oscillation and the feed-back will maintain this condition. There are applications for this kind of circuit and this is dealt with under "Oscillators".

Oscillators

Oscillators are of various kinds but they all depend upon the same principle for their operation. The essential condition for oscillation is positive feed-back from one circuit to another. If the feed-back is interrupted short bursts of oscillation will take place. If, however, it is continuous by virtue of the amplification of the valve the oscillations will be continuous. Provided the amplification of the valve is such that the excess of energy in the grid circuit is sufficient to overcome the losses in the valve, it will continue to oscillate.

The frequency of the oscillation is determined by the inductance and capacity of the tuned circuit. Fig. 26 shows a tuned grid oscillator. This is probably the most commonly used oscillator in receivers for normal broadcast purposes. One of its advantages is the ease with which oscillations can be obtained; another is that the component assembly offers no difficulty, and a particular advantage is that the tuned circuit is completely isolated from the power supply. The frequency of oscillation is determined by the condenser C and the

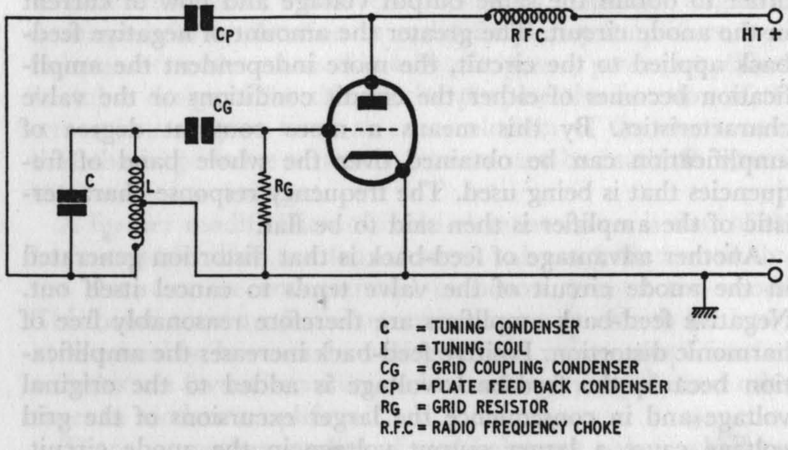


Fig. 26 Hartley tuned grid oscillator

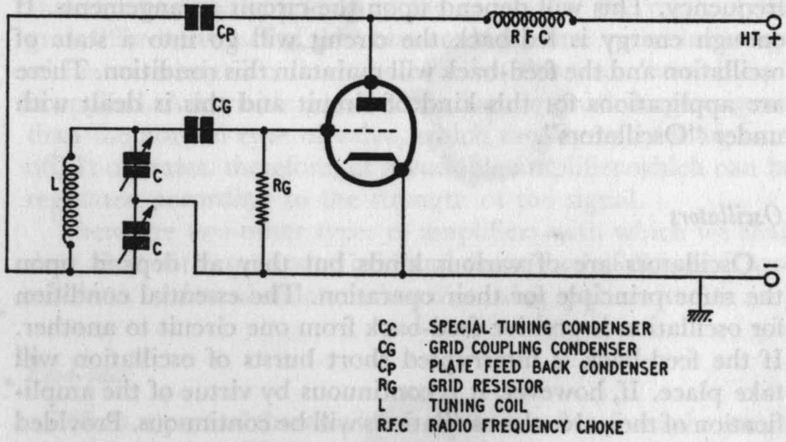


Fig. 27 Colpitts oscillator

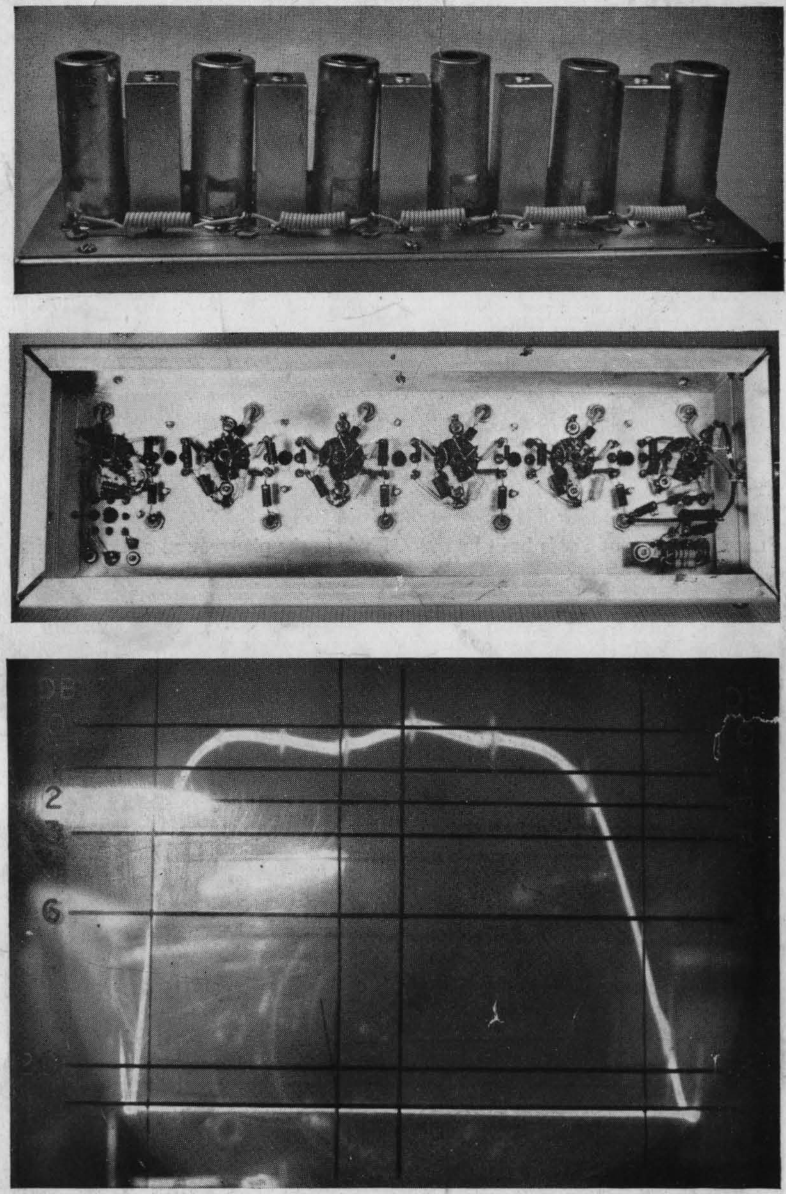


PLATE 3. A 30 mc/s I.F. amplifier built by R. I. Tomkins. The bottom illustration shows the 3 mc/s frequency response of the amplifier. (*Reproduced by courtesy of the B.A.A.*)

PLATE 4. Arrays of Yagis built by F. W. Hyde in the playground of the Salesian College, Battersea.

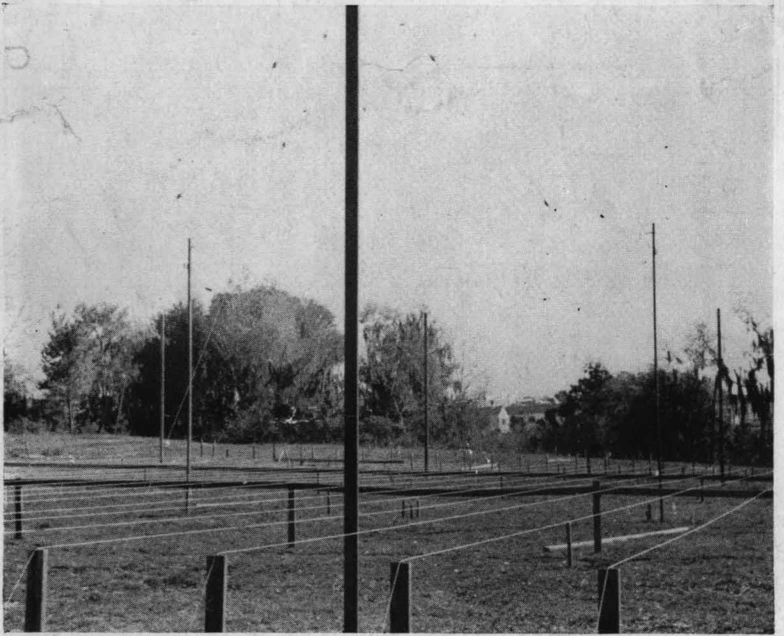


PLATE 5. A horizontal array of dipoles built by C. H. Barrow. The wire reflecting screen is clearly shown.

inductance L. The positive feed-back is obtained by the mutual coupling of the anode coil, L. This oscillator has a useful application up to a frequency of 50 megacycles. For higher frequencies, up to 300 megacycles, a more satisfactory arrangement is that of Colpitts. The Colpitts oscillator is illustrated in Fig. 27. It is a very convenient circuit to use and readily oscillates. One of the advantages of this circuit is that it lends itself to inductance tuning and can be kept reasonably stable.

Sometimes more than one valve is used in the process of producing oscillations. A typical instance of this is Fig. 28. This is called a "multi-vibrator". It will be observed that the anode of one valve is coupled to the grid of the other valve, so that energy is constantly being exchanged between these two valves and a state of oscillation exists. This is an oscillator which can be locked into step with an external control voltage, and later we shall see its application to certain techniques in radio astronomy.

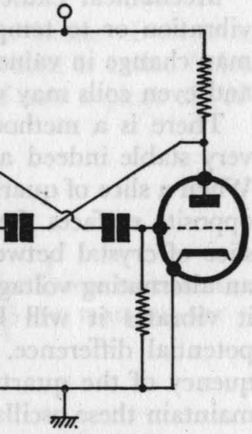


Fig. 28 Basic circuit of multi-vibrator

The oscillators that have already been described have used triode valves working as independent units. Sometimes an oscillator may be incorporated in a multi-electrode valve. Later, when dealing with receivers, we shall find that in the super-heterodyne receiver there is a stage which is called the "convertor" or "frequency changer". A valve used in this stage is usually a multi-electrode valve and part of its structure is devoted to the functions of an oscillator. One such type is the electron-coupled oscillator. The principle of this, as its name implies, is that the oscillating circuit is connected to the load only by means of the electron stream within the valve itself. By using it in this way changes in the load on the valve and the high-tension voltages have less effect on the actual frequency of oscillation. There is an improvement in frequency stability and the tuned circuit can readily be ganged with the other tuned circuits involved. One of the important requirements in an oscillator is that its frequency shall remain stable. Generally speaking, the causes of frequency variations will fall into two classes, electrical and mechanical.

Electrical variation can be due to changes in voltage in high tension supply, changes in valve characteristics during operation and random variations due to the condition of components in the circuit.

Mechanical causes of frequency variation can be due to vibration or to temperature. Components such as condensers may change in value under rising temperature in the receiver, and even coils may vary due to the expansion of the material.

There is a method by which we can render the oscillator very stable indeed and this is by the use of a quartz crystal. When a slice of quartz crystal has a potential applied across its opposite surfaces the crystal becomes strained. If we put this slice of crystal between the anodes of a condenser and apply an alternating voltage to it the crystal is caused to vibrate. As it vibrates it will be strained and develop an alternating potential difference. The result is always at the natural frequency of the quartz crystal. Once set into operation it will maintain these oscillations in an extremely stable manner over long periods. If we connect the crystal across either the anode or the grid coil of the oscillator, the crystal will control the generated frequencies. Certain types of converter used in radio astronomy require extremely stable oscillators. A crystal-controlled oscillator, therefore, provides the solution to the problem of stable oscillation.

An oscillator is, of course, the basis of the circuit for transmitting radio frequencies. Since in this case it is necessary that the frequencies shall have a high order of stability it is usual to have the transmitter controlled by an oscillator whose frequency is governed by a crystal.

Semi-conductors

Materials, such as germanium and silicon, are examples of semi-conductors. Their atomic structure is similar to that of insulators, but when a small amount of impurity is introduced during the manufacture of the crystal, it is possible for free electrons to exist and move through the crystals in the presence of an electrical field. It is possible for some of the atoms to be deficient in one electron, and as we have already seen, these electron deficiencies, or holes, move from atom to atom. Holes and electrons we call "charged carriers" when they are present in semi-conductors. One type of material is designated "P" type; this is because conduction takes place in the material by virtue of the deficiency in electrons. This is termed "hole conduction".

The other type of material having an excess of electrons operates by the method of conduction which is called "electronic". This is the "N" type material. If now a piece of "N" type material is joined to a piece of "P" type material and the voltage applied to them, a current will flow across the junction between the two and through the external circuit, see Fig. 29. The electrons indicated by the negative symbol are attracted across the junction from the "N" material, through the "P" material to the positive terminal of the battery. The holes indicated by the plus symbol will move in the opposite direction across the junction to the negative side of the battery. The current, therefore, flows through the circuit by reason of the electrons moving in one direction and the holes in the other direction. If we reverse the polarity of the battery, excess electrons in the material of "N" type are attracted away from the junction and the holes in the "P" material are attracted away from the junction by the negative potential of the battery.

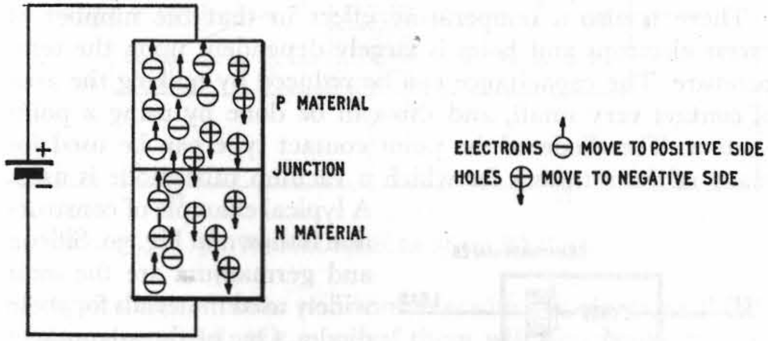


Fig. 29 Behaviour of a PN junction

The junction region is then left without any current carriers and there is no conduction.

The junction of the "P" and "N" type material acts as a diode or rectifier. There is, however, a measurable reverse current to the rectifier of this kind. In this respect it differs from the valve diode. The reason for the reverse current is that there are some carriers always left in the material; the carriers cross the junction by a process of diffusion and this takes place comparatively slowly. The fact that the junction forms a capacitor with two anodes of very small spacing implies that the capacitance is rather high. It is for this reason that there is a limit to the upper frequency at which such a semi-conductor will operate.

There is also a temperature effect in that the number of excess electrons and holes is largely dependent upon the temperature. The capacitance can be reduced by making the area of contact very small, and this can be done by using a point contact. The diode of the point contact type can be used for many of the purposes for which a vacuum tube diode is used. A typical example of construction is shown in Fig. 30. Silicon and germanium are the most widely used materials for these diodes. One of the advantages of the crystal diode is its small size, its low inter-electrode capacity and the fact that it requires no heater to release electrons.

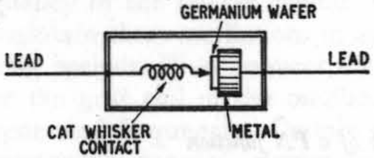


Fig. 30 Point contact diode

The junction type diodes of germanium and silicon are used generally as power rectifiers. There is a special form of silicon junction rectifier and these are known as "Zener" diodes. The characteristics of this diode allow it to be used somewhat as a voltage stabilizer, for it maintains a constant voltage drop over a wide range of current changes.

Another more recent type of diode is the tunnel-diode. This is a junction semi-conductor which exhibits negative resistance at certain voltages. This characteristic permits it to be used as an oscillator and as an amplifier. Oscillations have been obtained up to 2,000 megacycles and higher with this diode.

Transistors

If a sandwich is made of two layers of "P" type material with a thin layer of "N" type material in between, they are, in fact, two "PN" junction diodes back to back. If a positive voltage is applied to the "P" type material, marked "A" in Fig. 31, the current will flow through the left-hand junction, the holes moving to the right and the electrons from the "N" type material moving to the left. Of the holes moving into the "N" type material some will combine with the electrons there and be neutralized, but some of them will also travel to the right-hand junction. If the combination at "B" has a negative potential there will be no current flow in this circuit, but as there are now additional holes available at the junction to travel to this point and electrons can travel to point "A", a current can flow even though this section of the sandwich, considered alone, is in a condition to prevent conduction. Most of the current is between "A" and "B" and it does not flow out through the common connection to the "N" type material. This combination is called a "transistor".

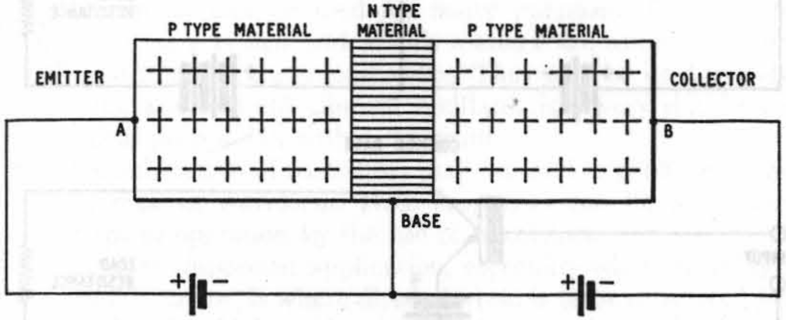


Fig. 31 Junction transistor of the PNP type

The three sections are known as the emitter, the base and the collector. The amplitude of the collector current depends upon the amplitude of the emitter current, so that the collector current is controlled by the emitter current. In practical types of transistors the emitter resistance is only of a low value, while the collector resistance is thousands of times greater. Power gains of the order of 20 to 40 db are therefore possible.

The transistor is made in several types. The "PNP" type has already been described, but the alternative arrangement of "NPN" may be used. The circuits are applicable to each of these two types.

Junction transistors usually have cut-off frequencies of the order of 50 megacycles. It is possible to increase the upper frequency limit by using a "drift" transistor. This has a particular distribution of the impurities in the base of the transistor which results in an electric field which accelerates the carriers across the junction. Transistors of this kind have cut-off frequencies as high as 100 megacycles.

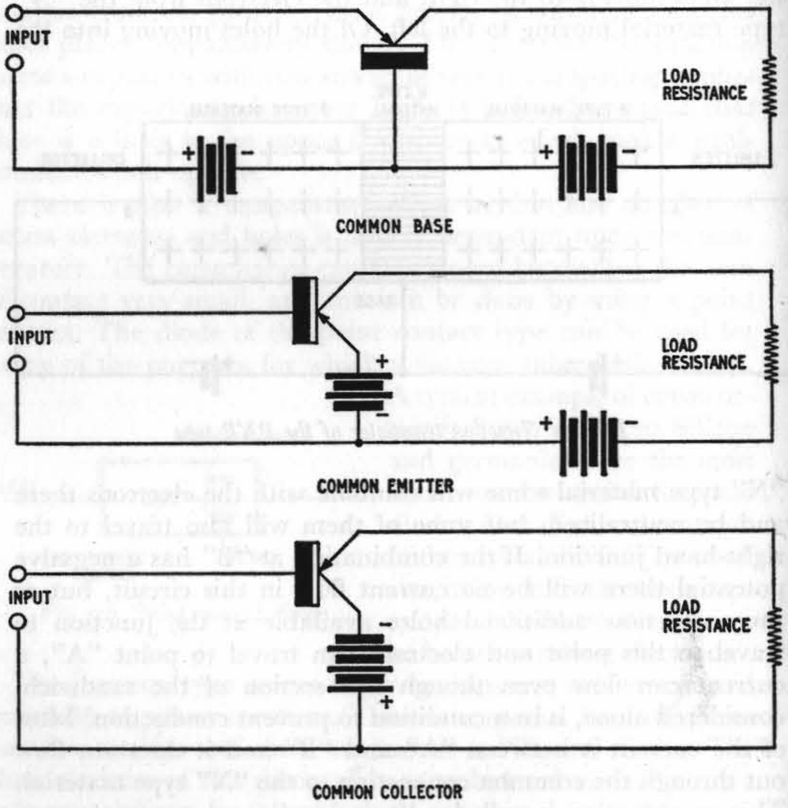


Fig. 32 Three basic amplifier circuits using transistors

A type of transistor known as "the surface barrier" type is available for operation at frequencies up to 250 megacycles and beyond. As its name implies, this uses an emitter and collector which are plated on a wafer of "N" type material.

Three basic transistor amplifier circuits are shown in Fig. 32.

Transistors may be used as oscillators and the frequency limit of oscillation is a function of the cut-off frequency. Providing there is sufficient feed-back, oscillation can be maintained just in the same way as it is with the valve oscillator.

Transistors can be used for many purposes. They can be operated as a switch and in this manner are somewhat more efficient than the thermionic valve. They may be used in pulse circuits, and a multi-vibrator oscillator is a very simple and practical proposition with a transistor.

A combination of transistors may be used to produce a saw-tooth type of waveform. D.C. amplifiers can be made very efficient in operation by the use of transistors.

Another important application, especially where stability is a primary factor, is where direct current is to be converted into A.C. signals which are proportional to the input. The name for this technique is "chopping". The use of a transistor here supersedes the former type of mechanical switch which used a vibrating reed.

Chapter Seven

AERIAL SYSTEMS FOR RADIO TELESCOPES

THE AERIAL IS the collector of energy and is the link between the free electro-magnetic waves in space and the receiver. In free space, and for our purposes we may regard the atmosphere as free space, the electro-magnetic waves travel with the velocity of light, that is 300 million metres per second, or 186,000 miles per second approximately.

If these waves meet with no barriers during their propagation they will normally tend to travel in straight lines. There are two components to an electro-magnetic wave – the electric field and the magnetic field.

In a previous chapter it was shown that if a potential was applied to a condenser an electric field was created between the condenser plates. In the case of an inductance the flow of current through the coils was shown to have created a magnetic field. Electro-magnetic waves have both these properties. The direction of the fields is at right-angles to the direction of the propagation of the wave. To designate this we use the term magnetic vector, when referring to the magnetic field, and the term electric vector when we are referring to the electric field. Polarization of electro-magnetic waves is said to be in the direction of the electric vector. In this condition the waves are plane or linearly polarized.

In radio astronomy we are concerned not only with plane waves but also with those that are polarized in other directions. Polarization may be right-handed or left-handed, and in both cases the waves are said to be elliptically polarized.

The ratio between the magnetic vector and the electric vector is always constant for a given medium in which propagation of electro-magnetic waves is taking place. We measure the electric vector by an arbitary value, E, in volts per metre and magnetic vector, H, we measure in terms of amperes per metre. This term can be represented in the more familiar manner as

ohms. There is a special value of this quantity for free space and it is called the "characteristic impedance of free space". This can be written in the form of the equation $\frac{E}{H} = 377$ ohms.

Energy is carried by electro-magnetic waves in the direction of propagation. It is possible to calculate the field strength that will be received at the aerial receiver if the strength of the source is known. Conversely, by measuring the amount of signal accepted by the receiver we can deduce the amount of power radiated by the source.

When dealing with aerials in the practical sphere we find it convenient to compare a new design with a standard half-wave dipole. A half-wave dipole is a convenient method of transferring the energy from the transmitter to free space, or collecting the energy from free space and conveying it to the receiver. The energy which is radiated or received by the aerial oscillates at high frequencies, that is radio frequencies. The smallest aerial that can be made to resonate at the frequency in use is a half-wave aerial.

The effectiveness of an aerial can be designated by its power gain. This is a quantity which depends upon the effective aperture of the aerial. If an aerial is placed in an electro-magnetic field which has a strength E, the power absorbed by the aerial will be

$$P = \frac{\lambda^2}{4 \times 377} \times 1{\cdot}635\ E^2 \text{ watts.}$$

This is dependent upon the effective aperture or area of the aerial. In certain cases the effective area may be that of the actual physical area of the aerial itself, though in others it may be less than the actual area occupied by the aerial. In Table 2 is shown a comparison of the type of aerial and its effective area.

From this table it will be observed that large broadside arrays of dipoles, reflector type aerials together with horn or lens aerials have an aperture equivalent to their physical size.

When we consider radio waves received from extra-terrestrial sources we describe them in certain terms related to the radiation received. This takes into account the spectrum, or range of frequencies, considered, their strength and the polarization

TABLE 2

TYPE	EFFECTIVE AREA
Half-wave dipole	$0{\cdot}13\lambda^2$
Half-wave dipole with reflector	$0{\cdot}25 - 0{\cdot}5\lambda^2$ depending on the spacing
Broadside array with reflector	Physical area
Broadside array without reflector	$0{\cdot}5 \times$ physical area

of the wave. The strength of the radiation can be described as the power which falls upon the Earth for a given area over a particular bandwidth. For radio astronomical purposes we can do this in two ways. Stated in MKS units, the flux density S is measured in watts per square metre per cycle per second (watts $M^{-2}(c/s)^{-1}$).

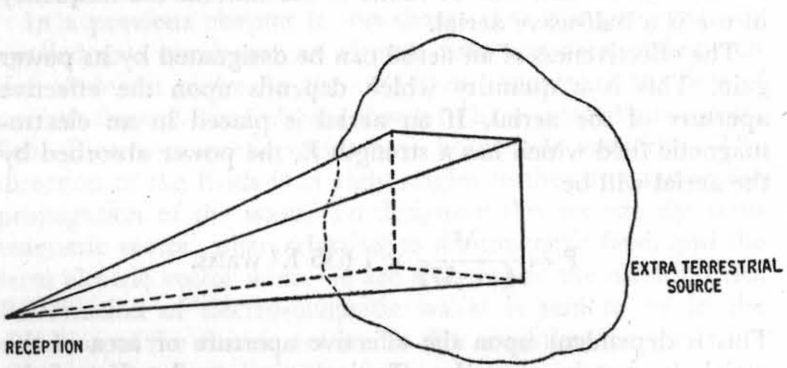


Fig. 33 Illustrating solid angle of reception pattern

For a distributed source of radiation we can speak of the brightness of a given area of the source of radiation. The flux density can also be described per unit of solid angle in watts per metre squared per cycle per second per steradian (watts $M^{-2}(c/s)^{-1}$ $ster^{-1}$). This is the usual form, and it is illustrated in Fig. 33.

An alternative system is based on Planck's law relating to

the emission of electro-magnetic energy at all frequencies from a black body, which is dependant upon the temperature of the body. Radio frequencies approximation to this law is given by the Rayleigh-Jeans formula:

$$B = \frac{2KT}{\lambda^2}.$$

This gives the brightness, B, in terms of the temperature T degrees Kelvin and the wavelength λ, where K is Boltzman's constant and is equal to $1{\cdot}38 \times 10^{-23}$ in MKS units. The brightness, B, may therefore, be specified terms of an equivalent black body temperature. It is convenient to use this formula to specify the strength of the received radiations when dealing with those sources emitting thermal radio waves.

When measuring the flux density with an aerial which is polarized in one direction it is necessary to remember that radiations from extra-terrestrial sources are randomly polarized and that therefore the average value of the flux is twice the total amount which impinges upon the aerial.

When describing the aerial of a radio telescope we say that it has an aperture of so many wavelengths at a specified frequency and it has a beamwidth of so many minutes or degrees of arc at the half-power points. Sometimes confusion is experienced by those who are new to electronics by the manner in which the terms "frequency" and "wavelength" are often substituted one for another in the same context. Dealing with radio telescopes the term "wavelength" will be used to describe either the size of a single aerial system or the distance between aerial systems, and the term "frequency" will indicate the particular point in the electro-magnetic spectrum where the observations are made.

Radiations are received at the aerial in the form of random voltages. The voltages are detected and amplified by the receiver and thence conveyed to the recorder, so the efficiency of the radio telescope aerial will depend very much upon the lowest level to which its internal noise can be reduced. Often the radiations which are received from extra-terrestrial sources are far below the noise level of the apparatus that we use to detect them. A low noise aerial and a low noise receiver is, therefore, an essential part of the equipment.

There are certain electrical characteristics associated with the aerial which are of great importance, they are:

(a) the impedance
(b) the directional properties
(c) the amount of forward gain
(d) the polarization
(e) the beamwidth.

All these properties are present, of course, in every aerial, but they have different relative values depending upon the frequency of operation and the purpose for which the aerial system is designed.

In Chapter I the limitation of the extent of the radio window in the spectrum was defined. In discussing the design of aerials, therefore, we will be limited to those which are applicable to radio astronomy and used within the limitations of the radio window. The design of the radio telescope aerial is determined by the observations to be carried out, but generally speaking we are concerned with aerials of good directional characteristics and well defined beamwidths with high gain. It is necessary that the aerial should be resonant, that is, it should be tunable exactly. This implies that there is a direct relation between the actual physical size of the aerial elements and the wavelengths in use. The ordinary broadcast receiving aerial using a long wire would not have very much application in radio astronomy.

The receiver must be carefully matched to the aerial, and the choice of aerial system will take into account the type of receiver that is to be used. Under normal broadcast conditions of working there is a considerable amount of tolerance in matching the aerial to the receiver, but since in radio astronomy we are dealing with extremely small voltages at the aerial, we require that the highest possible efficiency of the transfer of energy from the aerial to the receiver shall be achieved.

The directional property of an aerial enables it to differentiate between a wanted source of radiation and adjacent unwanted sources. The forward gain of an aerial is closely related to the directional properties and in general the more directional an aerial is the greater will be the forward gain.

A fundamental type of resonant aerial is the half-wave dipole. In its simplest form it consists of two wires or rods, each a quarter of a wavelength long and held rigidly by an insulator in the centre. The receiver is connected by means of a feeder at the two innermost ends of the wire or rod. In a half-wave dipole system this is the point of current maximum, as indicated in Fig. 34*a*. The impedance of this aerial is approximately 73 ohms.

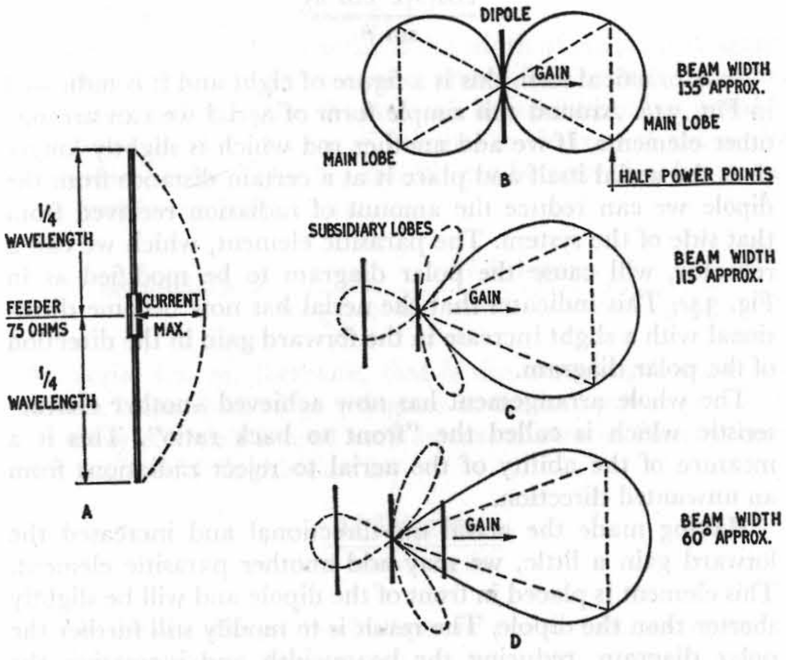


Fig. 34 Details of dipole and the effect of adding parasitic elements

The important characteristic of an aerial is the bandwidth. This is the efficient frequency range over which it will operate. In the dipole aerial the directional characteristics are substantially constant, but the impedance will change as the frequency is changed above or below the point of resonance of the aerial. The actual bandwidth of an aerial depends upon the size of the elements and their shape. A compromise is necessary between directional properties and the bandwidths, for a point can be reached where the bandwidth becomes too narrow for the purpose required. It is necessary in cases of this kind to sacrifice the forward gain in order to achieve the bandwidth required.

The directional characteristics of a dipole can be calculated because the aerial is symmetrical about its axis and its polar diagram will therefore be a circle. This will comply with the formula:

$$\frac{\cos(\pi/2\ \cos\ \theta)}{\sin\ \theta}$$

In a practical form this is a figure of eight and it is indicated in Fig. 34*b*. Around this simple form of aerial we can arrange other elements. If we add another rod which is slightly longer than the aerial itself and place it at a certain distance from the dipole we can reduce the amount of radiation received from that side of the system. The parasitic element, which we call a reflector, will cause the polar diagram to be modified as in Fig. 34*c*. This indicates that the aerial has now become directional with a slight increase in the forward gain in the direction of the polar diagram.

The whole arrangement has now achieved another characteristic which is called the "front to back ratio". This is a measure of the ability of the aerial to reject radiations from an unwanted direction.

Having made the aerial uni-directional and increased the forward gain a little, we may add another parasitic element. This element is placed in front of the dipole and will be slightly shorter than the dipole. The result is to modify still further the polar diagram, reducing the beamwidth and increasing the gain. This parasitic element is called a "director" and the effect on the polar diagram is shown in Fig. 34*d*.

It is possible to go on adding further directors in front of those already existing. Each one will get successively shorter in the forward direction. The practical limit of such directors is about thirteen. As each successive parasitic element is added there is a slight increase in forward gain, but a stage is reached where the forward gain no longer increases although the beamwidth has become extremely narrow.

For practical purposes there is an optimum overall length of the multi-element aerial and this is between 2·5 and 3 wave-

lengths in extent, measured from the reflector to the furthermost director. At high frequencies this type of aerial is extremely useful for it is simple to construct and of reasonably manageable size. Combinations of multi-element aerials may be used and when combined the overall gain is increased. There is also an improvement in the polar diagram which results in reduced beamwidth and with careful design a highly directional system can be obtained. A typical array of multi-element aerials is shown in Plate 4 and Plate 7.

The beamwidth of an aerial is the width of the polar diagram at the point where the received power falls to half its value. As stated earlier in the chapter, the beamwidth of an aerial is described as being so many minutes or degrees of arc at the half-power points. This beamwidth can be arrived at by means of a simple equation. It is:

$$\text{beamwidth} = \frac{57{\cdot}3\ \text{degrees}}{\text{length of the aerials in wavelengths}}.$$

An aerial system, therefore, that is one wavelength wide will have a beamwidth of 57·3 degrees at the half-power points. If an array of aerials is used such as that shown in Plate 7, then the beamwidth in this case is

$$\frac{57{\cdot}3\ \text{degrees}}{4\ \text{wavelengths}}.$$

The beamwidth at the half-power points will, therefore, be approximately 15 degrees.

The impedance of an aerial is determined by the system in which it is used. In the case of the multi-element aerials already described the position of the parasitic elements in the array has an immediate effect on the impedance of the dipole. With the reflector at an optimum distance of a quarter of a wavelength the effect is not sufficient to cause much mismatch. When, however, the director is added there is an appreciable reduction in the impedance of the dipole. In order to overcome this the dipole can be folded. This may be accomplished by a continuous length of tube being bent into shape as shown in Fig. 35*a*. An alternative method of constructing the dipole is to take two straight lengths of tube which are of differing diameters and joined together at the ends. The space between them and the relative diameters of the tubes control the impedance at the centre, Fig. 35*b*.

Another method is to use large diameter wire in the form of a loop. The configuration of the loop itself must be found by trial and error. Such a folded dipole has an impedance of approximately 250 ohms. The addition of the reflector and the first director reduce this to 75 or 80 ohms.

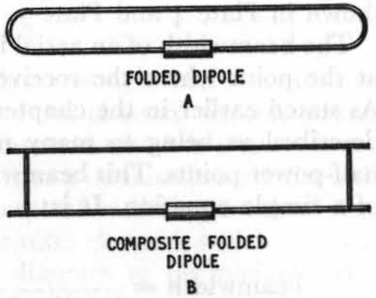


Fig. 35 Types of dipole

There are two other methods of achieving this object without folding the dipole. These two methods are called the "delta-match" and the "gamma-match". The delta-match uses a dipole which is continuous throughout its length in the form of a single rod. At equal distances from the centre a clip is fastened and two short pieces of tube are fixed to these clips at a distance of some two inches from the main dipole. To the inner side of the short tubes is connected the transmission line. The actual position of the taps on the dipole has to be determined by trial and error.

Impedance can be corrected using this method and so make the dipoles substantially independent of the rest of the array. An example of delta-match is shown in Fig. 36*a*. It is necessary to use a balanced transmission line with this method.

A gamma-match enables the same advantages to be obtained as with the delta-match and in addition allows the use of unbalanced coaxial cable as the link between the aerial and the receiver. The gamma-match, shown in Fig. 36*b*, again uses a dipole which is continuous throughout its length. The outer covering of the coaxial cable is connected to the centre of the dipole. The inner conductor of the coaxial cable is connected via a small capacitor to a point on one side of the centre of the dipole. The distance from the centre is determined by the impedance required. This type of dipole is often used in large fixed arrays where matching can then be done entirely by

coaxial cable. This avoids the use of balanced transmission lines.

Groups of multi-element arrays may be arranged in such a way that they can be steered in azimuth and altitude or equatorially. Generally speaking there is little advantage in the extra cost of the equatorial mounting since the best beamwidth that can be obtained is limited.

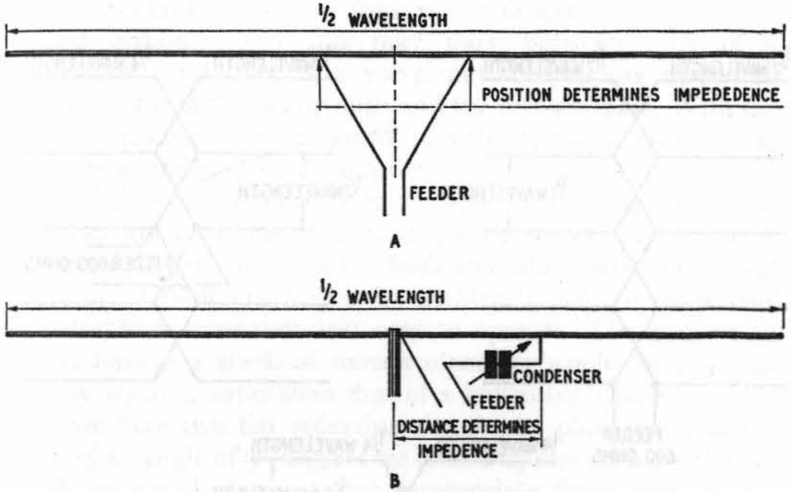


Fig. 36a & b Delta- and gamma-match (no insulator is required at centre)

There are, however, certain applications, such as that of satellite tracking, where an equatorial mount can increase the versatility of the system.

Reflector-type Aerials

Reflectors are used to modify the radiation pattern of an aerial. The backward radiation may be prevented from reaching an aerial by the use of a plain sheet reflector. If a suitably shaped reflector is used, control of the beamwidth is possible and the characteristics of the aerial system can be predetermined.

The arrangement of a large flat sheet reflector near a linear dipole aerial reduces the backward radiation. With small spaces between the aerial and the sheet a substantial gain can be obtained in the forward direction. The properties required of the sheet reflector can be preserved even when the size is reduced considerably. The extreme of this is, of course, the single element reflector behind a dipole. However, the sheet reflector does not come within the resonant range of the aerial itself and can therefore be used with a number of aerials covering a wide frequency range. The plain sheet can be replaced with an arrangement consisting of spaced wires or mesh. This may be used with a single dipole or a number of dipoles in the form of an array. A useful form of this is the Kooman array (Plate 4), which consists of a reflector made up of parallel wires in the same plane as the dipoles. The distance between the wires should not exceed one-eighth of a wavelength at the frequency being used. The reflector will then act substantially as a plain sheet. Alternatively, wire mesh can be used for the reflector, in which case the openings in the mesh should not exceed one-eighth of a wavelength.

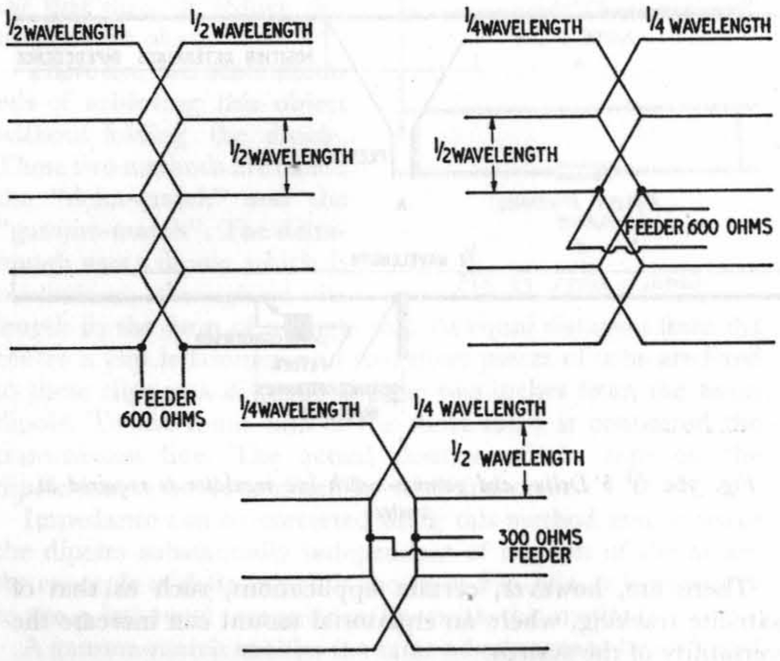


Fig. 37 Alternative connections for arrays of dipoles

Bearing these factors in mind, it will be possible to produce a reflector with wires spaced in a manner to suit the highest frequency at which it is proposed to operate. All lower frequencies will find the reflector increasingly efficient. Arrays of dipoles can be placed in front of the reflector at a predetermined distance, and this may vary between 0·15 of a wavelength and 0·25 of a wavelength depending upon the requirements of the beamwidth and the forward gain. With the smaller spacings there is usually greater forward gain, but at the expense of backward radiation.

A number of alternative arrangements of coupling the dipoles together are available and these are shown in Fig. 37. Such an aerial is very suitable for both steerable radiometers and fixed interferometers. From the amateur's point of view they are simple to construct and easy to operate. The gain of an aerial having a stack of sixteen elements would be approximately 19 db greater than that of a half-wave dipole.

If we take two flat reflecting sheets and place them in the form of an angle of 90 degrees we have a square corner reflector. With an aerial placed at the appropriate focal point within this we have an extremely efficient and highly directional aerial system. It is possible to use corner reflectors with angles which are greater or less than 90 degrees; however, 90 degrees gives the most efficient combination of directivity and efficiency. The spacing of the aerials from the apex of the reflector is extremely important because it has a marked effect on the impedance of the polar diagram. Fig. 38 shows the relative effects of spacing the aerial from the apex of the corner reflector.

At high frequencies, this is the type of aerial easy to construct and lends itself to a considerable range of frequencies for, provided the sides are made to an optimum length suitable for the lowest frequencies which will be used, aerials can be interchanged without disturbing the main assemblies. The aerials may be made fully steerable or, again, fixed in azimuth and steerable in altitude. Plates 8 and 9 show two such arrangements.

Corner reflectors need not be confined in size to that which accommodates one aerial, but can be made up long in respect to their height. The number of dipoles can then be spaced parallel with the reflector at the focal point and thus provide a means of increasing the resolution. Such a corner reflector will be "N" wavelengths long, and applying the formula for beamwidths, an aerial 4 wavelengths long would have a beamwidth of 15 degrees at the half-power points.

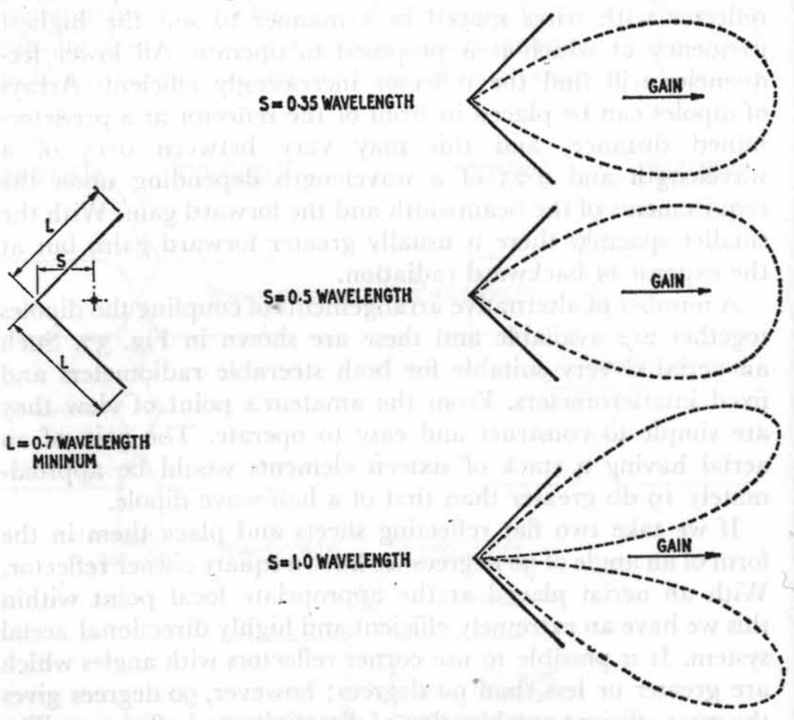


Fig. 38 Corner reflector and effect of distance of focal plane from apex

From the amateur's point of view this is one of the simplest forms of aerial for construction and full details of such aerial systems will be given in a later chapter.

Large fixed corner reflectors are used at the Mullard Observatories at Cambridge.

The next type of aerial to be considered is that of the cylindrical parabola. This is an improvement on the corner reflector, having greater efficiency by reason of its shape. The corner reflector differs in principle from the parabolic reflector, but for practical purposes the loss of efficiency due to the corner reflector is often outweighed by the extra cost of the cylindrical parabola.

Paraboloidal Reflectors

The principal advantage of the cylindrical parabola is that the maximum efficiency is achieved because the area of the parabola is fully illuminated and is in phase at all points. Fig. 39 shows a typical cylindrical paraboloid.

The principal requirements here are that the configuration of the shape of the paraboloid will be within one-eighth of a wavelength in contour on all points. At extremely high frequencies this can be a flat sheet or very close wire mesh, but at frequencies below 600 megacycles parallel wires are quite efficient. The large radio telescope at the Mullard Radio Observatories, designed by Professor Ryle, is an example of the cylindrical parabola. This has a reflector of parallel wires.

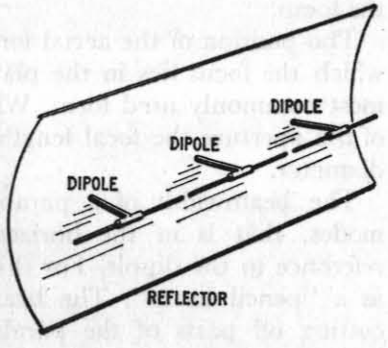


Fig. 39 Cylindrical paraboloid aerial

The construction of the cylindrical parabola is not beyond the skill of an amateur, since it can be fabricated in wood without any great difficulty. It is, however, expensive and unless the extra efficiency is fully justified several other types of aerials could be made for the cost of one cylindrical parabola.

The beamwidth of this aerial will be narrowest in the horizontal plane and widest in the vertical plane, in the same way as the corner reflector. For an aerial system ten wavelengths long the beamwidth at the horizontal would be of the order of 6 degrees.

If an aerial is placed at the focus of paraboloidal reflector and its electrical dimensions chosen such that the whole surface of the parabola is illuminated, radiations will be intercepted by the paraboloid and reflected as a plane wave of circular cross-section and may be collected by the aerial. The reflector surface of the paraboloid should not deviate from the true parabolic surfaces by more than an eighth of a wavelength at the highest frequency at which the aerial will be operating.

A paraboloid aerial offers certain very important advantages. Its main structure remains unaltered over the whole frequency range of its operation. In order to change frequency only the aerial assembly requires alteration. In most other types of aerials only polarization in the plane for which the aerial was designed can be used. With a paraboloid, however, all modes of polarization can be accepted, the only requirement being that the aerial of appropriate mode of polarization be placed at the focus.

The position of the aerial for maximum gain is the point at which the focus lies in the plane of the aperture. This is the most commonly used form. When the focus lies in the plane of the aperture the focal length is one quarter of the aperture diameter.

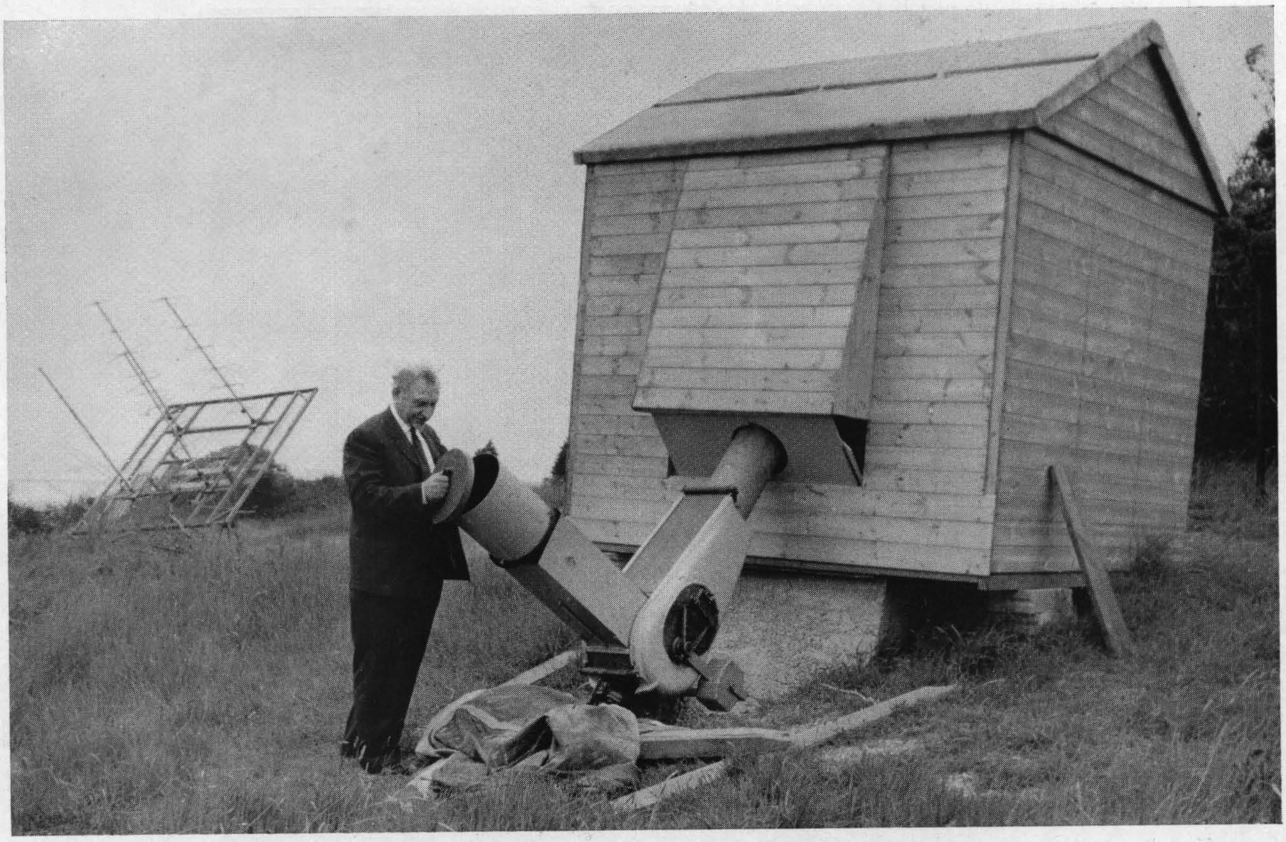
The beamwidth of a paraboloid is nearly equal in both modes, that is in the horizontal and vertical planes with reference to the dipole. For this reason it is usually spoken of as a "pencil beam". The beam-shape can be controlled by cutting off parts of the paraboloid. One of the paraboloid aerials used at Jodrell Bank has been designed in elliptical shape for the express purpose of shaping the beam.

Extremely high gains can be obtained with paraboloids when the diameter is large compared with the wavelength. If the aperture is uniformly illuminated, that is phase and amplitude are constant over the whole area of the aperture, the power gain of a paraboloid relative to a half-wave dipole is given by the following equation:

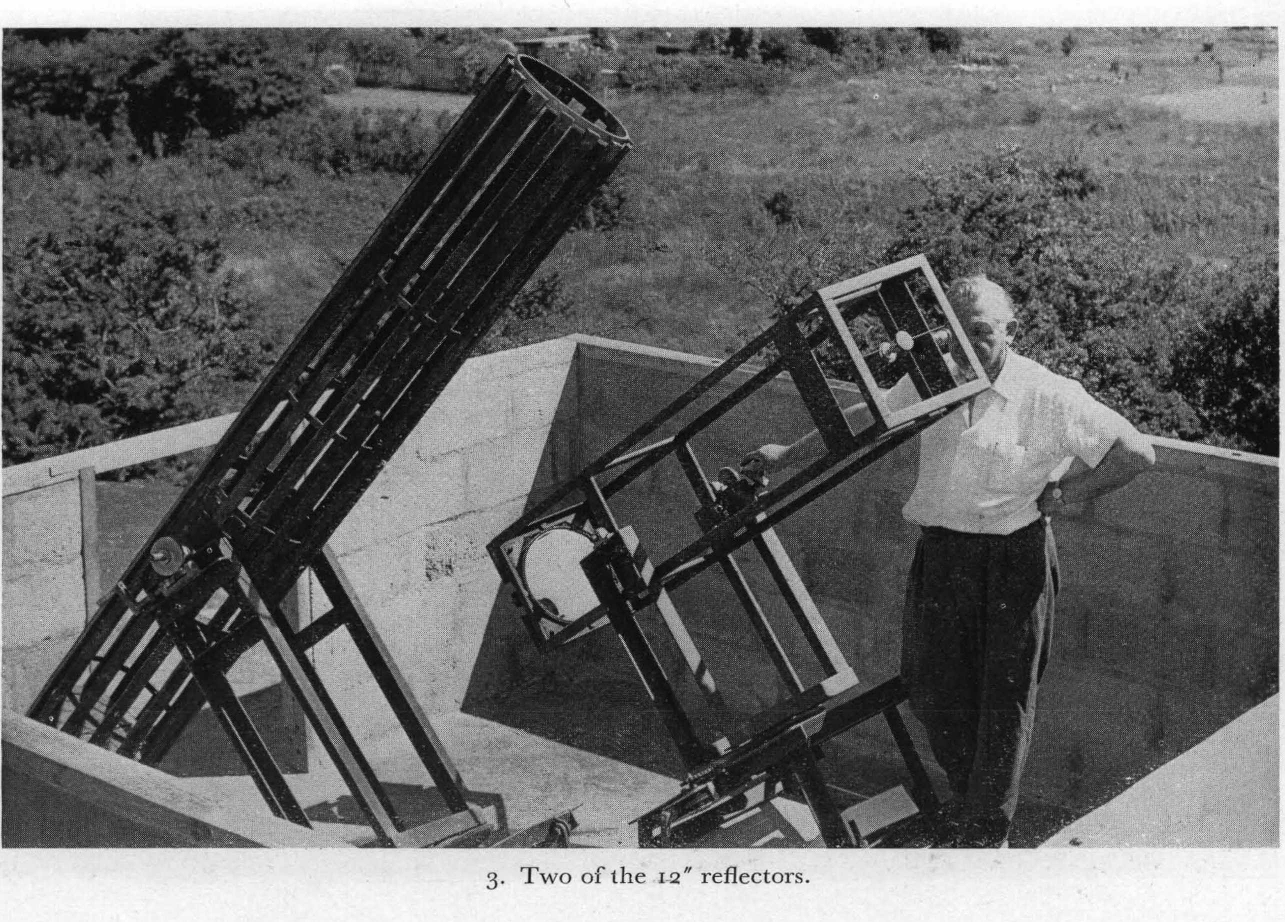
$$G = 6 \times \left(\frac{d}{\lambda}\right)^2.$$

If, at a certain wavelength, the paraboloid is twenty wavelengths in diameter it will have a gain of approximately 2,400 over a half-wave dipole. The polar diagram itself is a function of the ratio of the diameter to the wavelength. A narrow beam, which results from the use of a parabola, has small side lobes, the largest of which is not more than ten or twelve per cent of the main beam. This means that an extremely high degree of

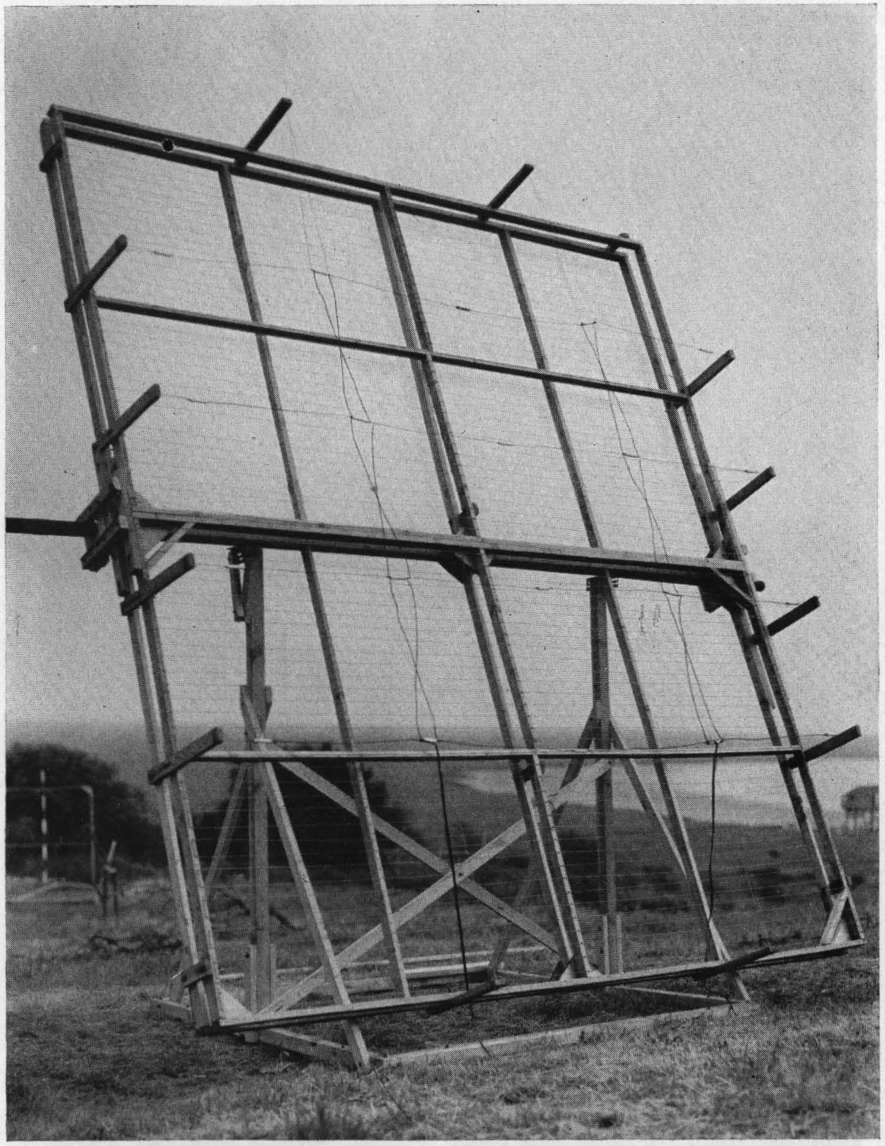
1. The author's observatory in a Martello Tower at Beacon Hill, St. Osyth, Essex



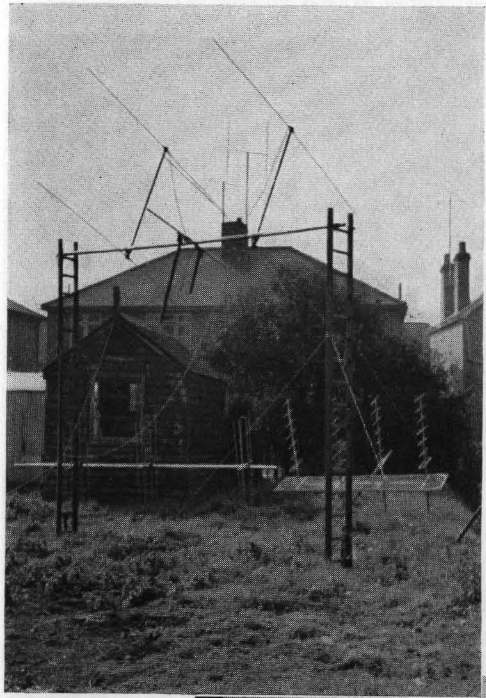
2. 9″ Coudé refractor with the 240 megacycle radiometer in the background.



3. Two of the 12″ reflectors.

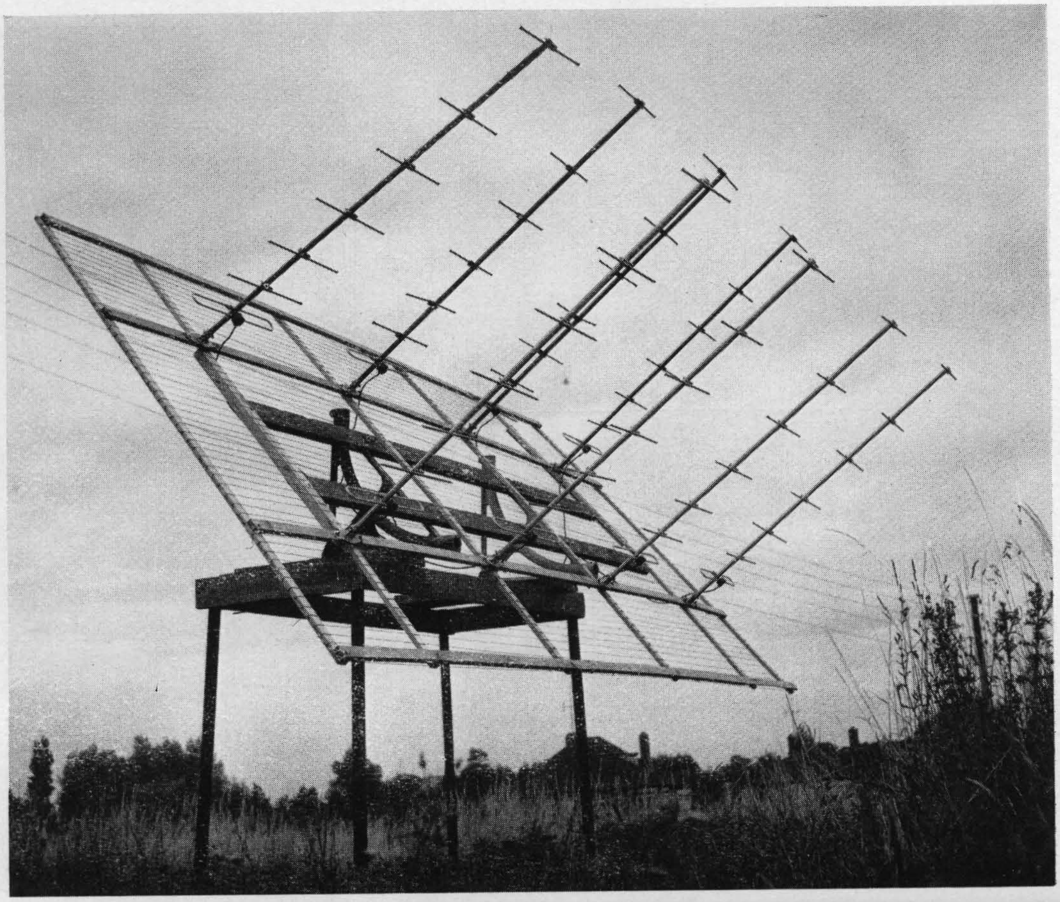


4. Kooman array.



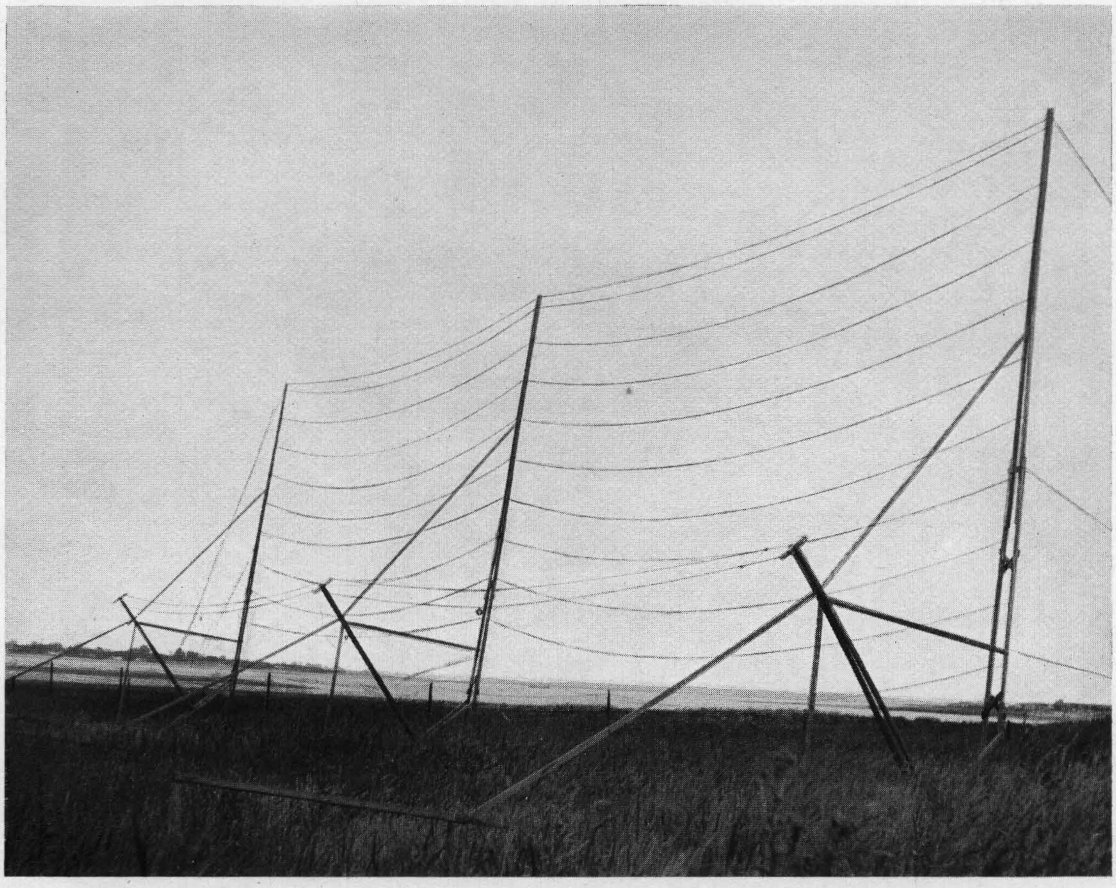
5. (*left*) One of the author's early aerials erected in the back garden.

(*below*) Observing with a spectroscope fitted to the Coudé refractor.

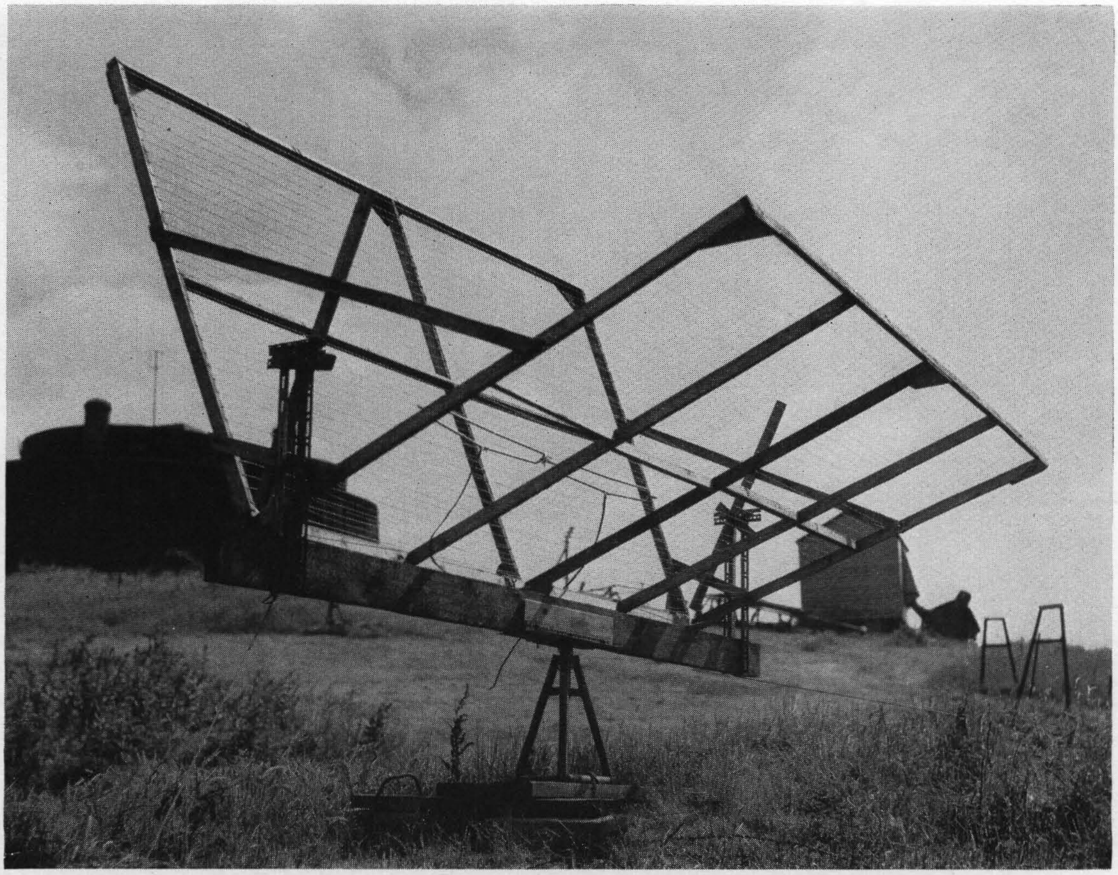


6. Multi-element aerial.

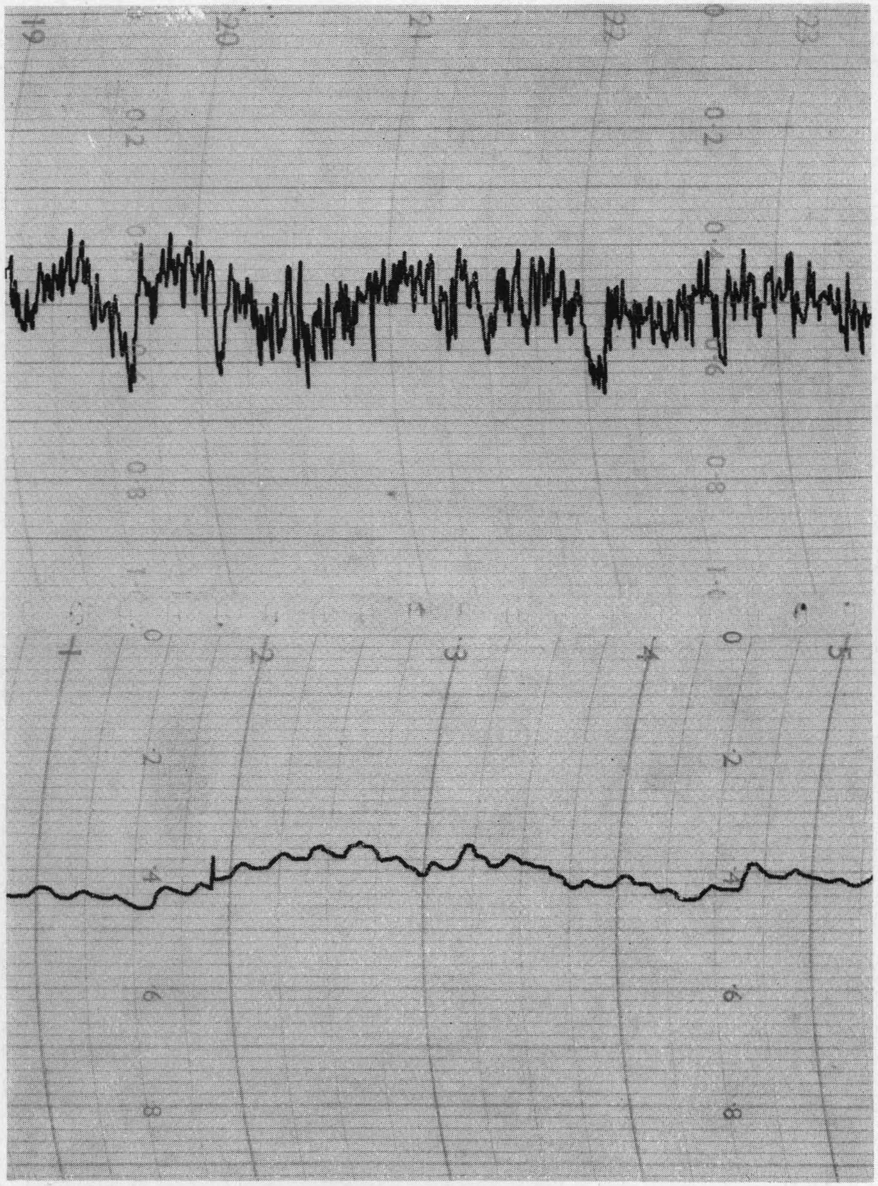
7. Small corner reflector for 220 megacycles.



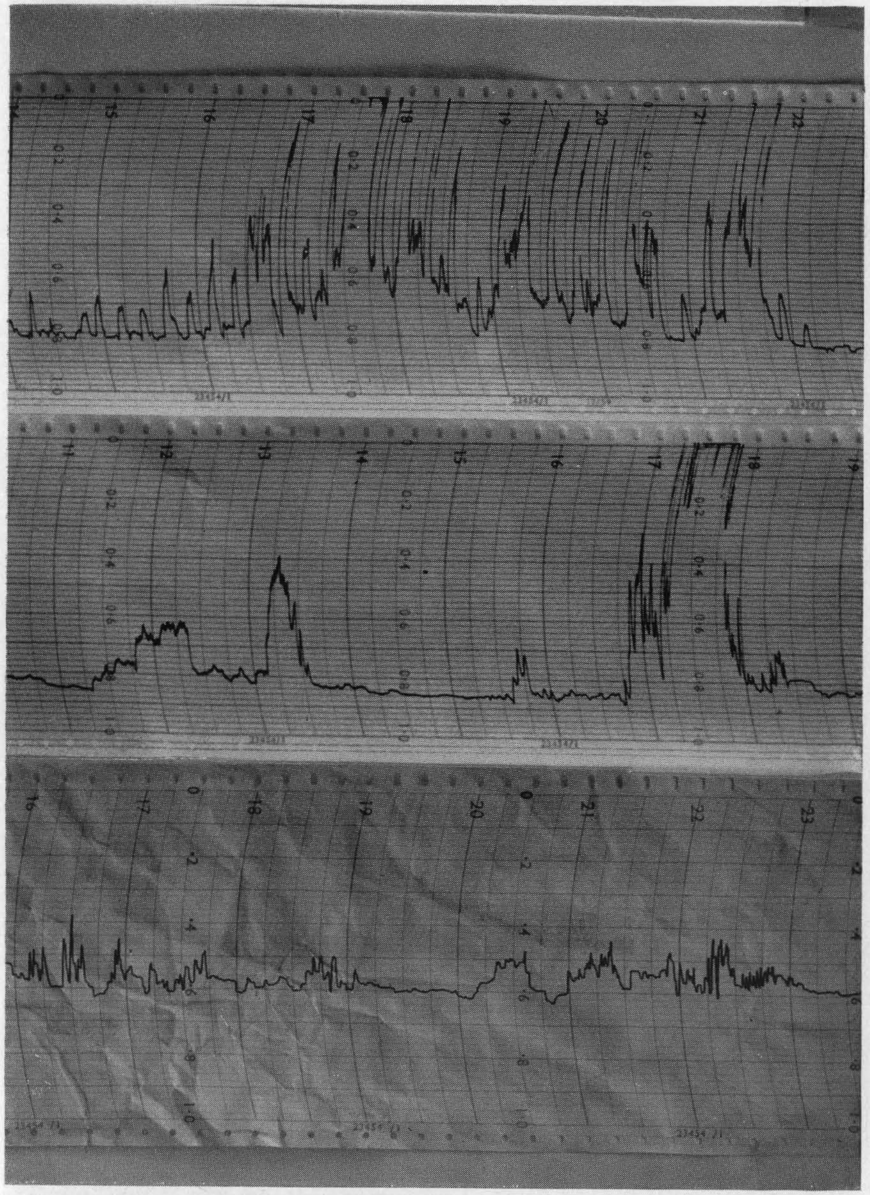
8. Large corner reflector for 26–27 megacycles.



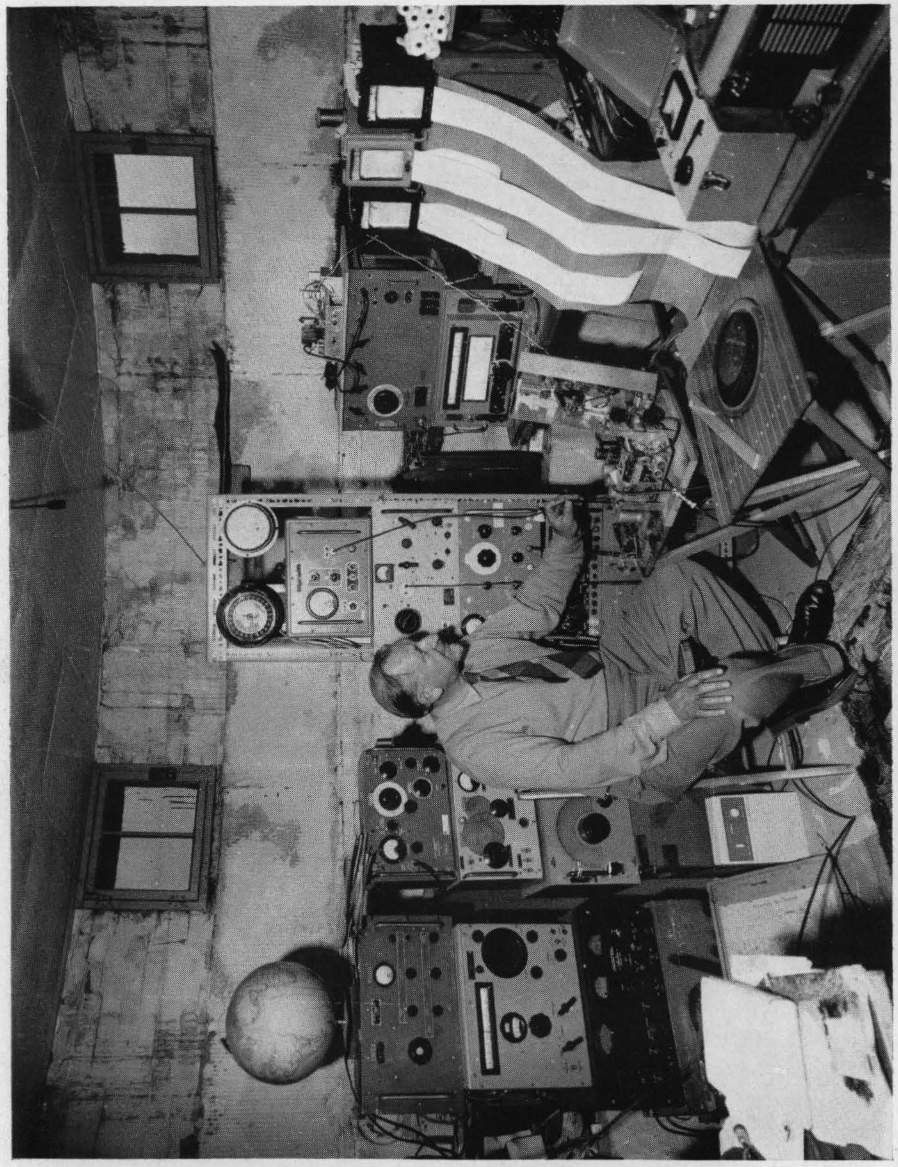
9. Steerable corner V.



10. Comparison of time constants
Upper trace: short time constant 1–2 secs.
Lower trace: long time constant 6–8 secs.



11. Typical traces of (*top*) The Sun passing through the aerial beam; (*centre*) Part of the Milky Way (Sagittarius Area); (*bottom*) Two radio sources which show the effects of scintillation.



12\. The control room which is called the eye-piece of the radio telescope. The author constructed some of this himself, and the remainder is surplus Government equipment suitably modified.

rejection of unwanted radiations is achieved. In practice it is never possible to get the illumination of the bowl entirely uniform as it usually tapers off towards the outside edge.

A certain amount of control of distribution can be achieved by the method of feed adopted. If a dipole with a reflector is used in order to suppress radiation falling directly in line with the aerial, beamwidth is increased by about twenty-five per cent. The increase in beamwidth will, of course, result in some decrease in gain. This is usually about thirty per cent. The compromise will depend upon the purpose for which the aerial is to be used.

Helical Aerials

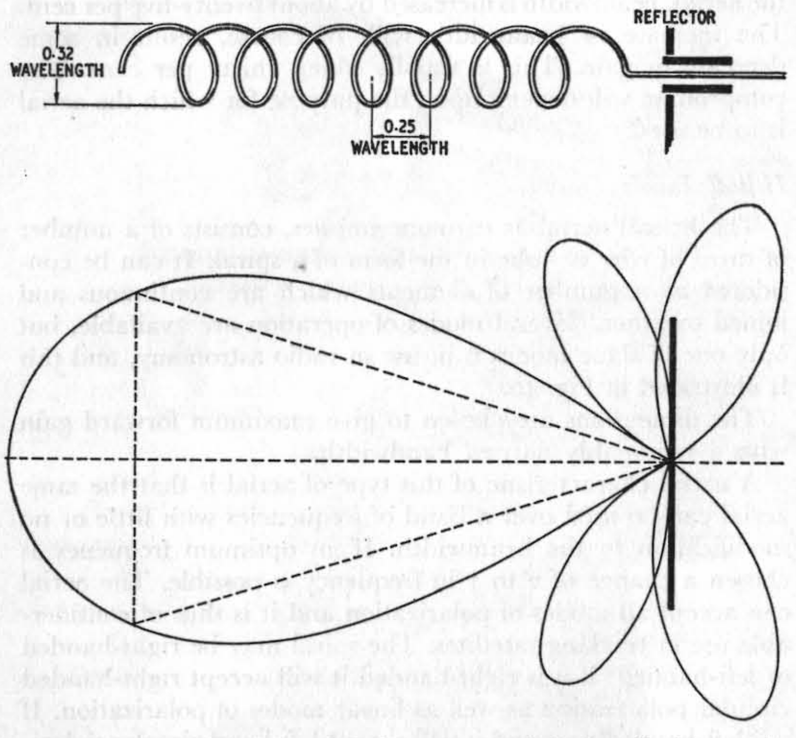
The helical aerial as its name implies, consists of a number of turns of wire or tube in the form of a spiral. It can be considered as a number of elements which are continuous and joined together. Several modes of operation are available, but only one of these modes is in use in radio astronomy, and this is illustrated in Fig. 40.

The dimensions are chosen to give maximum forward gain with a reasonably narrow bandwidth.

A useful characteristic of this type of aerial is that the same aerial can be used over a band of frequencies with little or no modification to the beamwidth. If an optimum frequency is chosen a change of 2 to 1 in frequency is possible. The aerial can accept all modes of polarization and it is thus of considerable use in tracking satellites. The spiral may be right-handed or left-handed; if it is right-handed it will accept right-handed circular polarization as well as linear modes of polarization. If it is left-handedly wound it will accept left-hand circular polarization and all modes of linear polarization.

The dimensions of the aerial vary considerably and are dependent upon the purpose which for the aerial is to be used. Helices may be used singly or in groups; a typical example of the use of the helix in groups is at the Italian observatory at Arcetri. Here a number of helices, each covering its own frequency band, are mounted on a sheet reflector. By this means a very wide range of frequencies are covered in regular observations of the Sun. Typical dimensions of a single helix are shown in Fig. 40, and the resultant polar diagram in Fig. 41.

The reflector used with a helix must be at least a half-wavelength in diameter. Its impedance at the termination point is approximately 80 ohms, it can therefore be coupled to the receiver by ordinary coaxial cable. Generally speaking increasing the length of the helix and the number of turns will decrease the beamwidth; doubling the axial length of the helix would reduce the beam by some thirty per cent.



Figs. 40 & 41 Helical aerial of 7 turns at a pitch of 120° with a total length of 1·6 wavelength and its resultant polar diagram

Rhombic Aerials

A rhombic aerial has certain special applications. As its name implies it is rhombic in shape; it may be erected horizontally above the ground with its elements parallel to the ground, or it may be mounted on an assembly which raises it above the ground upon a steerable platform. The advantages offered by the rhombic aerial are that this, like the helix, can operate over a frequency range of 2 to 1. For certain applications in radio astronomy such an aerial is extremely useful. It was the basis of one form of spectro-heliograph used in Australia by J. P. Wild.

A number of rhombics, each arranged to cover a specific band of frequencies, were mounted in such a manner that they could follow the Sun. A special receiver attached to them enabled a band of frequencies to be swept from 40 to 240 by means of a rotating tuning device. With this instrument the astronomers were able to follow the progress of flares. A typical rhombic aerial mounted above ground is shown in Fig. 42.

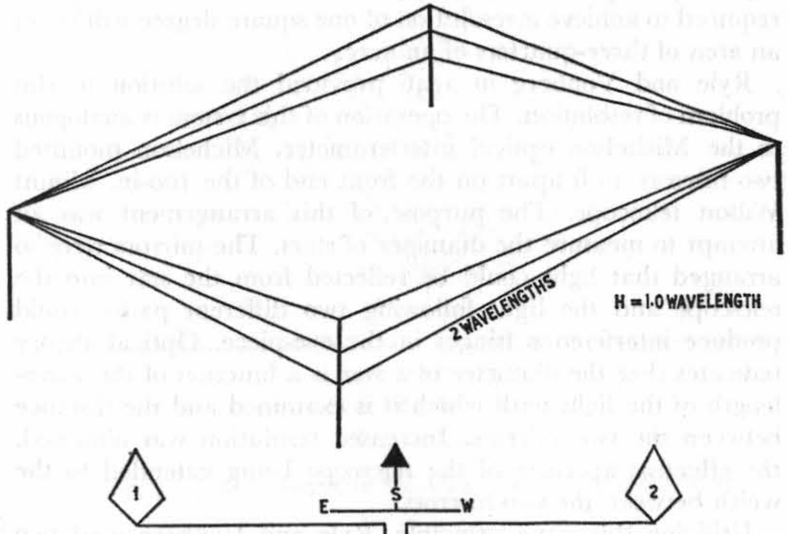


Fig. 42 Rhombic aerial with typical dimensions

A rhombic aerial is sometimes used at the low frequency end of the radio spectrum, with sides two or more wavelengths long, but, the aerial becomes somewhat extensive in size. Except at high frequencies this aerial will offer little interest to the amateur, for better resolution and narrower beamwidths can be obtained by other methods.

Resolution of Aerials Systems

The problems of high resolution with radio telescopes has been mentioned previously. With single aerial systems high resolving powers can only be obtained when the wavelengths in use are small compared with the aperture of the telescope. With telescopes having large parabolic reflectors, like that of Jodrell Bank, the degree of resolution that can be obtained at the centimetre wavelengths is very high. However, when working in the region of metre wavelengths resolution is greatly reduced. This is a problem of physical size. For example, at a frequency of 200 megacycles the size of the aerial system required to achieve a resolution of one square degree will cover an area of three-quarters of an acre.

Ryle and Vonberg in 1946 provided the solution to this problem of resolution. The operation of this system is analogous to the Michelson optical interferometer. Michelson mounted two mirrors 20 ft apart on the front end of the 100-in. Mount Wilson telescope. The purpose of this arrangement was an attempt to measure the diameter of stars. The mirrors were so arranged that light could be reflected from the star into the telescope and the light following two different paths would produce interference fringes in the eye-piece. Optical theory indicates that the diameter of a star is a function of the wavelength of the light with which it is examined and the distance between the two mirrors. Increased resolution was obtained, the effective aperture of the telescope being extended to the width between the two mirrors.

Utilizing this same principle, Ryle and Vonberg used two separate aerial systems on a long base-line. The energy collected by the two aerials was fed together into a receiver. With such an arrangement, the aerial system as a whole has a polar diagram in the shape of a fan beam, in which occur a number of interference patterns or lobes. The width of each lobe will be dependent upon the angular separation between the points of minimum signal. The angular separation decreases as the distance between the aerials of the system is increased, so if the

aerials are separated by twenty wavelengths the minimum points of the lobes will be separated by approximately 3 degrees. A further separation of the two aerials will reduce the width between these minimum points. If now a radio source is observed and this source is of smaller diameter than the width of separation of the minimum points, the output of the receiver will vary in a periodic manner. For a source which is greater in extent than the width of the lobes, the output of the receiver will maintain a fairly constant level. From this it will be appreciated that the resolution of the telescope is increased enormously. In Fig. 43 the diagram shows the arrangement of the aerials and the consequent pattern.

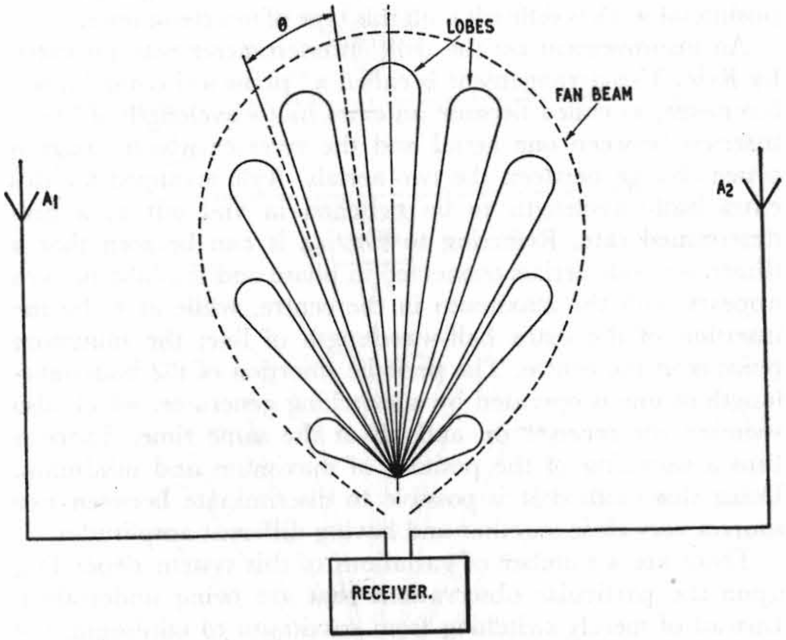


Fig. 43 Lobe formation in fan beam of a 2-aerial interferometer

The two aerials are usually fixed on an east to west base-line and the rotation of the Earth serves to sweep the beam across the sky. Each individual aerial system can be moved in altitude and therefore successive strips of the sky may be scanned.

Though valuable work can be done with this simple type of interferometer, it has certain disadvantages. If an extended source of radiations is near to a smaller source, it may be difficult to resolve the two, the larger source giving a continuous level, and this level will only be slightly varied by the smaller source variations superimposed upon it. Unless the exact position of the lobes is known without any ambiguity, exact positional work is difficult with this type of interferometer.

An improvement on the drift interferometer was provided by Ryle. This arrangement is called a "phase switching" interferometer, so called because an extra half-wavelength of line is inserted between one aerial and the receiver which causes a phase change between the two aerials. Ryle arranged for this extra half-wavelength to be switched in and out at a predetermined rate. Referring to Fig. 44 it can be seen that *a* illustrates both aerials connected in phase and the lobe pattern appears with the maximum in the centre, while at *b*, by the insertion of the extra half-wavelength of line, the minimum point is in the centre. The periodic insertion of the half-wavelength of line is operated by a switching generator, which also switches the receiver on and off at the same time. There is thus a sampling of the position of maximum and minimum. Using this method it is possible to discriminate between two sources very close together and having different amplitudes.

There are a number of variations of this system depending upon the particular observations that are being undertaken. Instead of merely switching from maximum to minimum, the whole lobe system may be made to rotate. This arrangement is therefore called "rotating lobe interferometer".

Another type of pencil beam radio telescope that gives high resolution is the Mills Cross. The lines of dipoles are arranged in corner reflectors in the form of a large cross. Very high resolution was obtained with this aerial system. The beamwidth was reduced to some 50 minutes of arc. Each arm of this cross was approximately 1,500 ft along and each array consisted of some 250 half-wave dipole elements. Each of the individual arms had a fan-shaped polar diagram covering some 50 degrees by 0·6 of a degree. The pencil beam response was obtained by

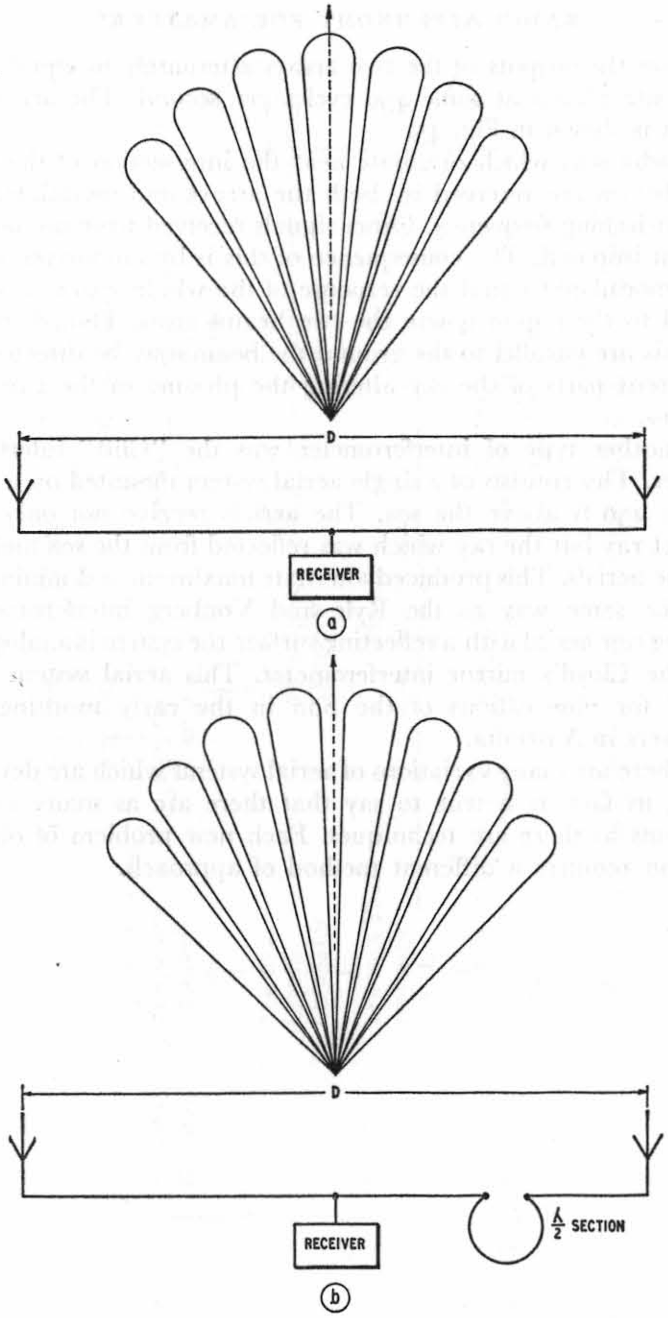


Fig. 44a & b Phase switching interferometer reception pattern mixing the outputs of the two arrays alternately in equal and opposite phases at some 450 cycles per second. The arrangement is shown in Fig. 45.

Radiations which originate near the intersection of the two fan beams are received by both the arrays and modulated at the switching frequency. Other signals received have no modulation imposed. The consequence of this is by amplifying only the modulated signal the response of the whole system is confined to the region where the two beams cross. Though these aerials are parallel to the ground the beam may be directed to different parts of the sky altering the phasing of the two fan beams.

Another type of interferometer was the "Cliff" interferometer. This consists of a single aerial system mounted on a cliff some 250 ft above the sea. The aerials receive not only the direct ray but the ray which was reflected from the sea and up to the aerials. This produced alternate maximum and minimum in the same way as the Ryle and Vonberg interferometer. Using one aerial with a reflecting surface the system is analogous to the Lloyd's mirror interferometer. This aerial system was used for observations of the Sun in the early morning by workers in Australia.

There are many variations of aerial systems which are devised and, in fact, it is true to say that there are as many aerial systems as there are techniques. Each new problem of observation requires a different method of approach.

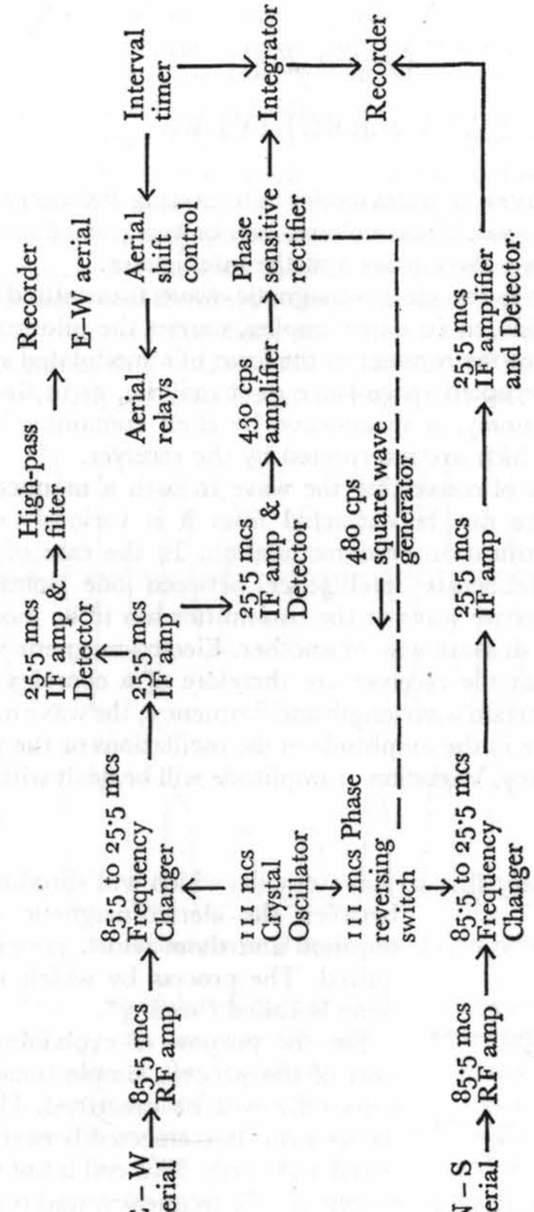


Fig. 45 Equipment layout of special receiver for use with Mills Cross aerial system

Chapter Eight

RADIO RECEIVERS

A RADIO RECEIVER is a device for interpreting voltages which arrive at the aerial. These voltages can be synthesized into two principal parts – the carrier and the intelligence.

The carrier is the electro-magnetic wave transmitted from the source which, as its name implies, carries the information. It may arrive at the receiver in the form of a modulated space-wave, an interrupted space-wave or it may be, as in the case of radio astronomy, a space-wave by itself containing many components which are interpreted by the receiver.

The process of converting the wave in such a manner that the intelligence can be extracted from it is variously called detection, rectification or demodulation. In the case of communicating deliberate intelligence between one point and another the carrier wave at the transmitter has to be modified or modulated in some way or another. Electro-magnetic waves which arrive at the receiver are therefore of a complex type. They have a certain wavelength and frequency; the wave may be modified either in the amplitude of the oscillations or the variation of frequency. Variation of amplitude will be dealt with first.

Detection

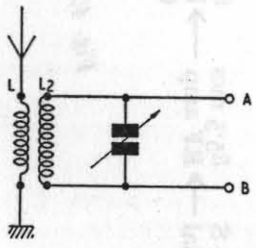


Fig. 46 Simple tuned circuit

It is necessary first to have circuits which will discriminate between the electro-magnetic waves required and those which are not required. The process by which this is done is called "tuning".

For the purpose of explaining this part of the process, simple tuned circuits only will be described. Fig. 46 shows a coil L, connected between the aerial and earth. This coil is not tuned to any specific frequency, and is therefore called "aperiodic". It is loosely coupled to another winding, across which is placed a condenser. We have now a radio frequency transformer, and voltages from the aerial will be applied across the coil L and transferred to the coil L_2. As we have already seen, when a coil and a condenser are balanced a condition of resonance arises in which the circuit becomes selective and highly responsive to a particular frequency of oscillation. Across the points A and B, therefore, there will appear a large voltage.

We now have a complex wave at a point where we can begin to discriminate between the intelligence carried by the electro-magnetic wave and the electro-magnetic wave itself, the carrier.

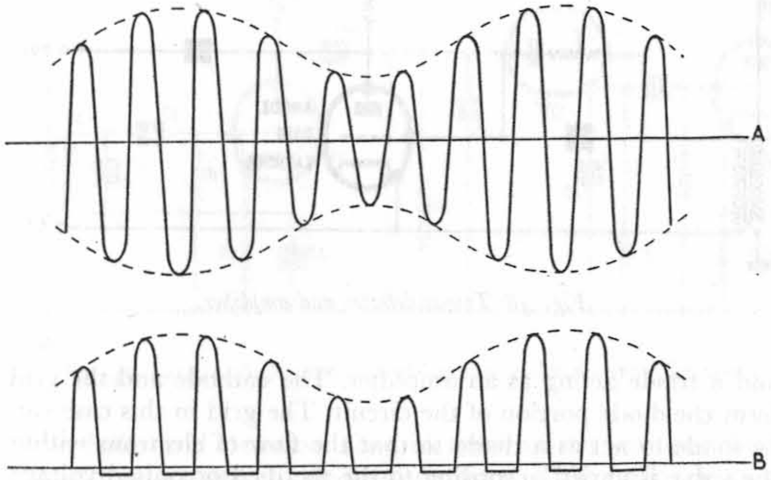


Fig. 47 Modulated carrier A before and B after rectification

The operation of the diode was shown to have the quality of rectification; that is, it was able to accept the flow of current in one direction only. Referring to Fig. 47, it will be seen that at A there is a modulated carrier. This is the varying voltage which will appear between the ends of the coil and transferred by means of the transformer to the points A and B. If we now connect a diode, which may be a crystal or a valve, across these two points in series with a condenser and resistance combination, we shall reduce the modulated carrier to the condition shown in Fig. 47*b*. We now have a pulsating current flowing through the resistor and the passage of this current will cause a voltage to appear across the resistor. If we substitute a pair of telephones for the resistor then the variations of current, that is the pulses, will become audible in the receiver. We have thus completed the process of detection, rectification or demodulation.

This is the simplest form of receiver. The diode is not the only form of detection; it is comparatively insensitive and needs a fairly high voltage to operate it. In the simple form of receiver described, only a very powerful signal would produce any results in the telephones. We can, however, make use of a triode valve to perform the operation of detection. This circuit, which is shown in Fig. 48, is really a combination of a diode rectifier

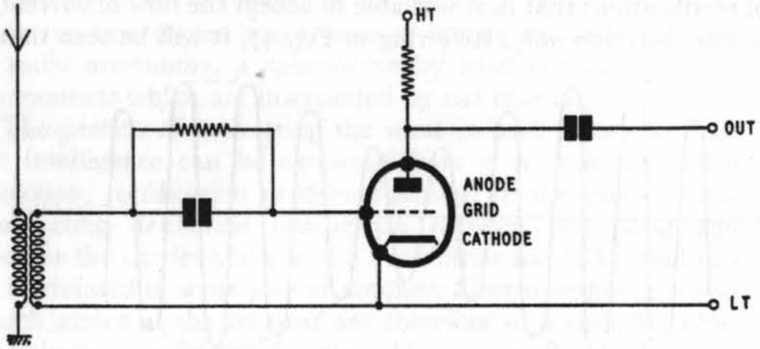


Fig. 48 Triode detector and amplifier

and a triode acting as an amplifier. The cathode and the grid form the diode portion of the circuit. The grid in this case can be made to act as a diode so that the flow of electrons within the valve is varied according to the rectified or pulsed voltage which appears across the small condenser and resistor. The variation in anode current will be in sympathy with the intelligence which has been carried on the electro-magnetic wave. This type of detector is known as the "leaky-grid" detector. It has certain disadvantages in operation in that the incoming intelligence is not completely separated from the radio frequency and consequently distortion is introduced. This is not important if coded signals are being received, but if it is intended for speech or music it can have certain disadvantages such as distortion of the original intelligence.

The output from this detector may be further amplified by following it with a low frequency amplifier. Such a simple amplifier has already been described and is shown in Fig. 22.

The sensitivity of a simple receiver can be increased by feeding back some of the voltage in the anode circuit and adding it to the voltage in the grid circuit, as was decribed previously in the section on "feed-back" (pages 104–5). If the amount of feed-back is controlled a very great increase in sensitivity can be obtained. This is called "reaction" and was a very popular form of receiver in the early days of broadcasting.

There are other types of detector apart from the diode and

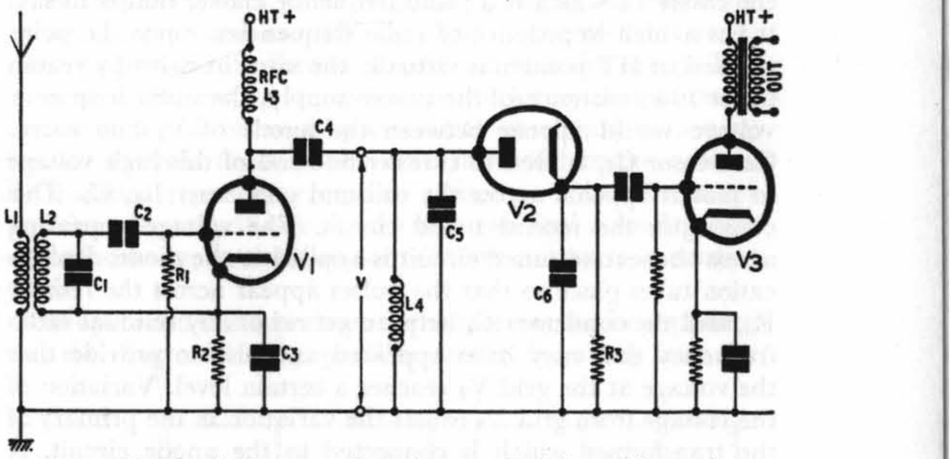


Fig. 49 Tuned radio frequency, diode detector and triode amplifier receiver

the leaky-grid detector. One of these which was popular at one time was the anode-bend detector. This made use of the curvature of the valve's characteristic and was introduced as a form of detector with low distortion. It was an attempt to improve on the leaky-grid detector, but the use of improved circuits, with the high gains obtainable with other types of receiver where a diode can be used, render this circuit obsolete.

Tuned Radio Frequency Receivers

By placing a valve or number of valves, in front of the diode tuned at radio frequencies, a much higher voltage can be applied to the diode. Fig. 49 shows a tuned radio frequency receiver. The voltage appears across the coil L_1 and is conveyed to coil L_2 in the usual manner. It is then amplified by the valve and an amplified voltage appears across L_4 and has all the components of the original electro-magnetic wave. This much higher voltage is available for detection by the diode which then provides a rectified output across R_3 which is applied to the amplifier.

The actual operation of the circuit is as follows: The aerial coupling circuit consists of the coils L_1, L_2 and C_1; this is the tuned circuit. After amplification the voltage appears across the choke L_3 which is a radio frequency choke, that is to say, it has a high impedance of radio frequencies. Since the point marked at HT positive is virtually the same as earth by reason of the low resistance of the power supply, the radio frequency voltage would appear between the anode of V_1 and earth. Condenser C_4, is used to transfer one end of this high voltage so that it appears across the coil and condenser L_4, C_5. This constitutes the second tuned circuit. The voltage appearing across the second tuned circuit is applied to the diode. Rectification takes place so that the pulses appear across the resistor R_3, and the condenser C_6 helps to get rid of any residual radio frequency that may have appeared and also to provide that the voltage at the grid V_3 reaches a certain level. Variation of the voltage from grid V_3 causes the variation in the primary of the transformer which is connected to the anode circuit. A loudspeaker connected to the secondary circuit, and an output transformer, provides the means of making the intelligence carried by the electro-magnetic wave audible to the listener.

Superheterodyne Receivers

The most common type of radio receiver in general use nowadays is the superheterodyne. The superheterodyne principle differs from those circuits which have already been described. It consists of:

(a) a tuned circuit and a signal frequency amplifying stage (in some very simple receivers the signal frequency stage is omitted)

(b) mixer or frequency changer
(c) intermediate frequency amplifier
(d) detector
(e) output stage.

The incoming signal is passed by the aerial circuits to the signal frequency amplifier where it is amplified and passed on to the mixer. Here a local oscillator signal is mixed with the amplified signal frequency.

The purpose of this mixer stage is to change the signal frequency from one level to another and this is usually from a higher frequency to a lower frequency. The lower frequency enables much greater amplification and more selective circuits to be designed for the further amplification of the signal. This chain of amplifiers, or intermediate frequency amplifiers as they are called, is followed by a detector which may be a valve diode or a crystal diode or any other form of detection system. The rectified signal may now be passed to a low frequency amplifier, the output circuit of which can feed a loudspeaker or telephones or, in the case of radio astronomy, a recording system.

The operation of a superheterodyne receiver can be explained from the diagram in Fig. 50. The carrier frequency from the transmitter is applied to the aerial circuits which may then be amplified at the signal frequency by a signal frequency amplifier, the output of which is fed into the mixer or converter. This, as its name implies, is the point where the incoming signal is mixed with a local signal and converted to the intermediate frequency which is then amplified by the intermediate amplifier

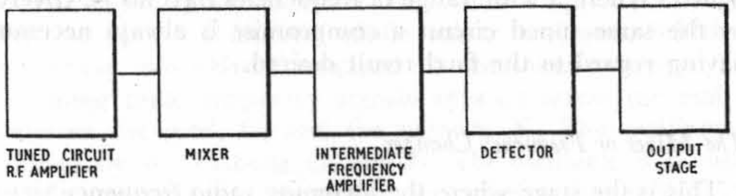


Fig. 50 Block diagram of superheterodyne receiver

stages. This is then followed by the detector stage and then the low frequency stage and finally to the loudspeaker or recorder.

For the sake of regularity we will deal with this diagram and describe the processes step by step. The input tuned circuit we have already dealt with under simple receivers. The signal frequency circuit we have mentioned and a simple form of this was shown in Fig. 46. The signal frequency, or radio frequency stage, amplifies the elected incoming signal. These amplifiers are, of course, all operated at radio frequencies, but the tendency has grown up for this to be called "signal frequency" when discussing superheterodyne receivers in order to distinguish between the amplification before the mixer and the second stage of amplification which is at the intermediate or lower frequency.

Design of signal frequency circuits is important; they must provide an adequate bandwidth if all the intelligence conveyed by the incoming signal is to be passed on. In the case of an ordinary radio receiver it will be necessary to provide tuned circuits which will give a substantially even amplification over the band of frequencies for which the receiver is designed. The narrower the bandwidth required, the greater can be the degree of efficiency obtained with such a circuit. If this band is comparatively narrow, extremely efficient stages can be designed. Where a wide range of frequencies have to be covered by the same tuned circuit a compromise is always necessary having regard to the final result desired.

The Mixer or Frequency Changer

This is the stage where the incoming radio frequency signal is mixed with that of the local oscillator. The combining of these two provides in the output of the frequency changer the original radio frequency signal modulated by the local oscillator. Present in the anode circuit of the frequency changer is the original carrier frequency plus the difference frequency between that of the oscillator and the carrier frequency, and also another component equal to the carrier frequency minus the oscillator frequency. The tuned circuit in the anode of the mixer is arranged to respond to one of these difference frequencies.

This new frequency, which will contain all the components of the modulation carried by the original incoming signal, can now be amplified more efficiently. The name for this stage is now apparent, it converts the high radio frequency incoming signal to a much lower frequency signal which can then be amplified without difficulty. It can be described as a mixer because two frequencies are mixed together, and it can also be described as a frequency changer, because it changes the high frequency incoming signal to a lower frequency. The actual difference frequency will depend largely upon the purpose for which the receiver is to be used. For broadcast purposes the differences between the oscillator frequency and the incoming frequency is usually 465 kilocycles.

If we now take a frequency in the broadcast band of 1,000 kilocycles, which is 300 metres wavelength, the frequencies present in the frequency changer stage will be the 1,000 kilocycles from the incoming signal and the 1,465 kilocycles of the oscillator. The difference frequencies will be 465 kilocycles. The intermediate amplifying system which follows is tuned to the difference frequency. This being a relatively low frequency enables circuits to be arranged with much greater efficiency than the high frequency circuits.

There are a number of forms of frequency changers and Fig. 51 shows three possible versions corresponding to three particular types of valves when used as a frequency changer. Fig. 52 shows a hexode triode mixer oscillator, which is essentially two valves contained in the same envelope. The incoming radio frequency signals appears across the coil L_1 between the grid, G, and the cathode K. The oscillator is formed by the cathode G_2 and P_1. The oscillator is a tuned anode-type oscillator, the tuning circuit being that at TO. The oscillator is injected into the grid system G_2. A radio frequency signal which is already modulating the electron stream between anode 2 and the RF grid is now further modulated by the signal injected to G. The difference frequencies will appear in the anode circuit and of these we select the required intermediate frequency by the tuned circuit TIF. Fig. 51*a* shows a pentode triode mixer and the operation of this will be clear from the diagram. Fig. 51*b* illustrates another combination of valve which is an octode. Here the first three electrodes serve

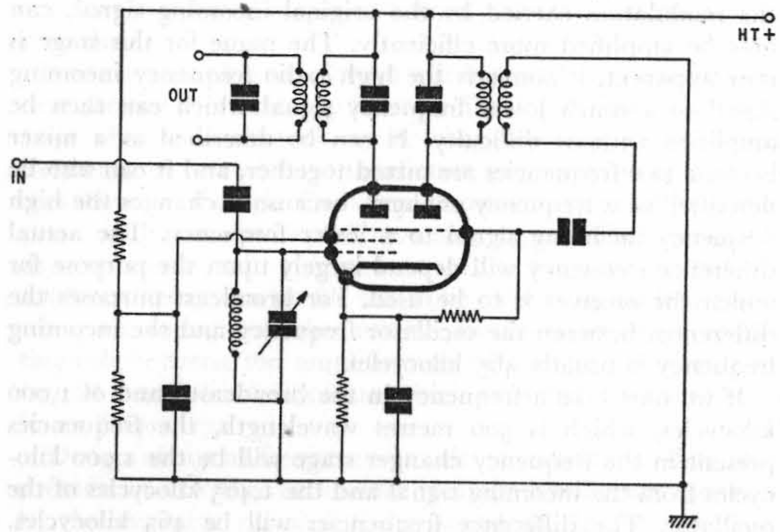


Fig. 51a Pentode triode mixer

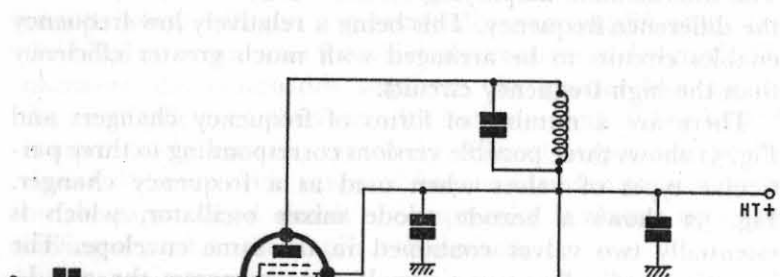


Fig. 51b Octode mixer

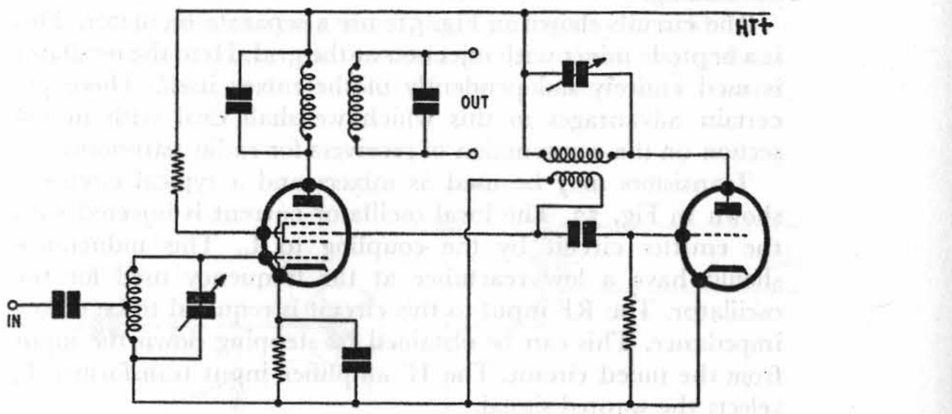


Fig. 51c Heptode mixer with separate triode oscillator

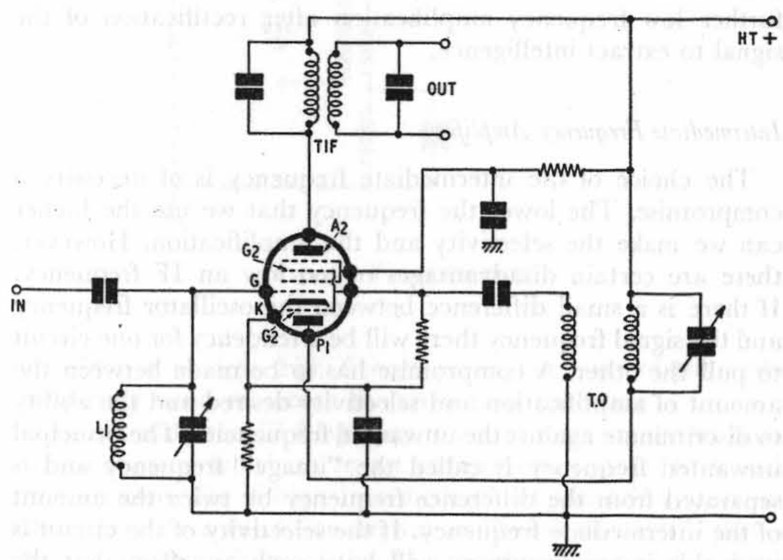


Fig. 52 Hexode triode mixer

for the oscillator; the remainder is for the incoming signal and the mixing.

The circuits shown in Fig. 51*c* use a separate oscillator. This is a heptode mixer with injection at the grid. Here the oscillator is used entirely independently of the mixer itself. There are certain advantages in this which we shall deal with in the section on the construction of receivers for radio astronomy.

Transistors may be used as mixers and a typical circuit is shown in Fig. 53. The local oscillator current is injected with the emitter circuit by the coupling to L. This inductance should have a low reactance at the frequency used for the oscillator. The RF input to this circuit is required to be of low impedance. This can be obtained by stepping down the input from the tuned circuit. The IF amplifier input transformer T, selects the wanted signal.

The signal can now be amplified by one or more stages of intermediate amplification.

The intermediate amplifying stages are followed by normal methods of detection and the detector stage can be followed by further low frequency amplification after rectification of the signal to extract intelligence.

Intermediate Frequency Amplifier

The choice of the intermediate frequency is of necessity a compromise. The lower the frequency that we use the higher can we make the selectivity and the amplification. However, there are certain disadvantages of too low an IF frequency. If there is a small difference between the oscillator frequency and the signal frequency there will be a tendency for one circuit to pull the other. A compromise has to be made between the amount of amplification and selectivity desired and the ability to discriminate against the unwanted frequencies. The principal unwanted frequency is called the "image" frequency and is separated from the difference frequency by twice the amount of the intermediate frequency. If the selectivity of the circuit is poor this image frequency will have such an effect that the same signal may be tunable in two places on the tuning scale. The efficiency of a converter is its ability to provide a low image signal. Up to a frequency of 7 megacycles the 465 kc intermediate frequency gives a satisfactory performance both from the point of view of the oscillator pulling and the image ratio. For frequencies greater than 7 megacycles and up to about 50 megacycles an intermediate frequency of 1·5 megacycles is satisfactory. In general the intermediate frequency should not be less than five to seven per cent of the radio frequency band on which the receiver is operating.

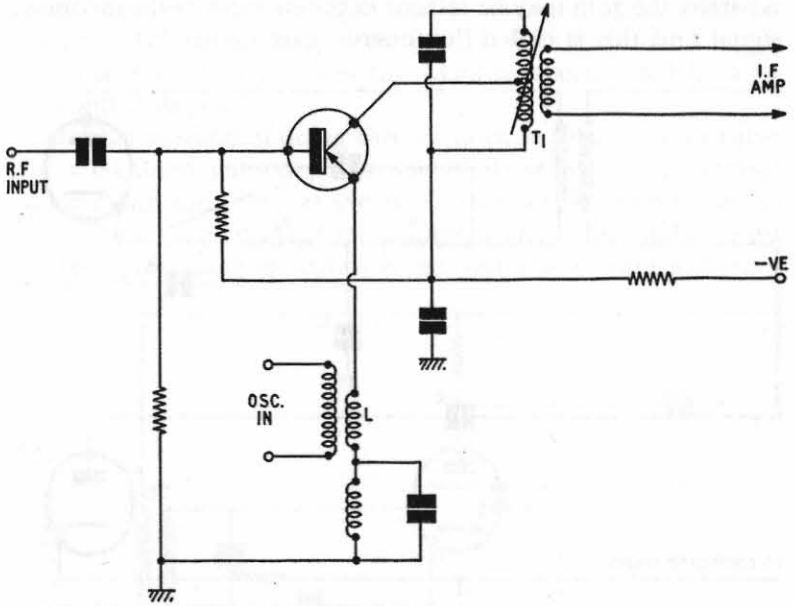


Fig. 53 Transistor mixer

In choosing the IF frequency it is necessary to avoid those frequencies on which there are commercial services or broadcast stations, for if high intermediate amplification is in use there may be a condition known as "break through", that is to say, the sensitivity of the IF amplifier being high, signals may be sufficiently strong for these to be received into the IF circuits and detected as though this were in itself the complete receiver.

Automatic Gain Control

With highly sensitive receivers it is necessary to have regulation of the gain of the receiver, for all signals received are not of the same strength. Wide ranges of variation would produce a change in modulation which, depending upon the setting of the gain control, may be below, or far above, the threshold level of audibility. Some system, therefore, is necessary whereby the gain may be related to the strength of the incoming signal and this is called "automatic gain control".

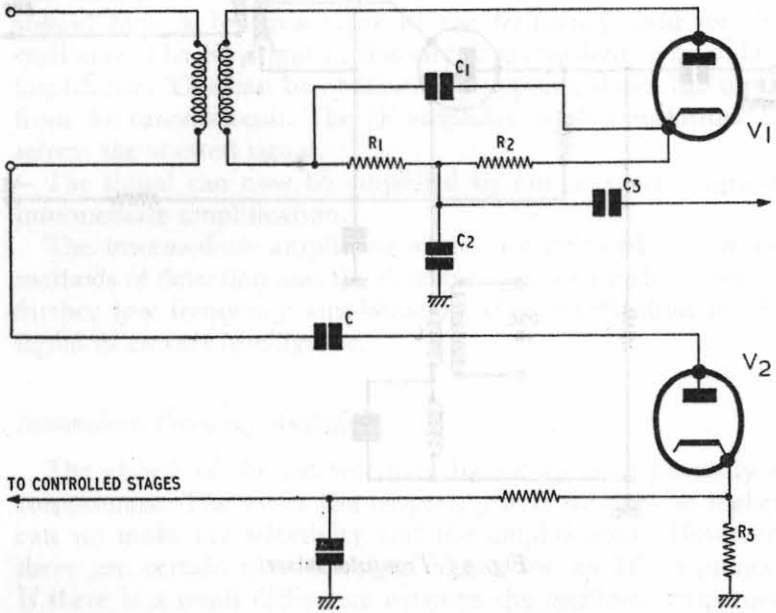


Fig. 54 AGC using separate diode valves

A rectified D.C. voltage which is developed by the receiver signal across the resistor in the detector circuit can be used to vary the bias on the radio frequency and the intermediate amplifier valves. This voltage will be proportional to the average amplitude of the signal and the gain can be reduced automatically as the signal strength increases. Control of at least two stages is necessary in a normal receiver. Although one method of deriving the AGC voltage is from the load of the diode detector, it is better to use a separate AGC rectifier.

A typical circuit is shown in Fig. 54. V_1 is the diode detector and the signal developed across R_1 and R_2 is coupled to the audio stages through C_3 and C_1. The resistor R_1 and C_2 are to filter out the unwanted radio frequency signal which remains. The AGC valve V_2 takes a small proportion of the signal via C from the last IF transformer and the rectified voltage is developed across R_3. V_1 does not rectify on weak signals because the fixed bias that is arranged at R_3 must be exceeded before the valve can rectify. The negative voltage which is developed provides the AGC bias and is fed back to the control stages.

Another method of doing this is shown in Fig. 55. This valve has a common cathode, and one diode serves as the normal rectifier and the other as the AGC rectifier. It is necessary to introduce a time constant for the operation of the AGC circuit where rapid fading of signals is encountered a rapid following

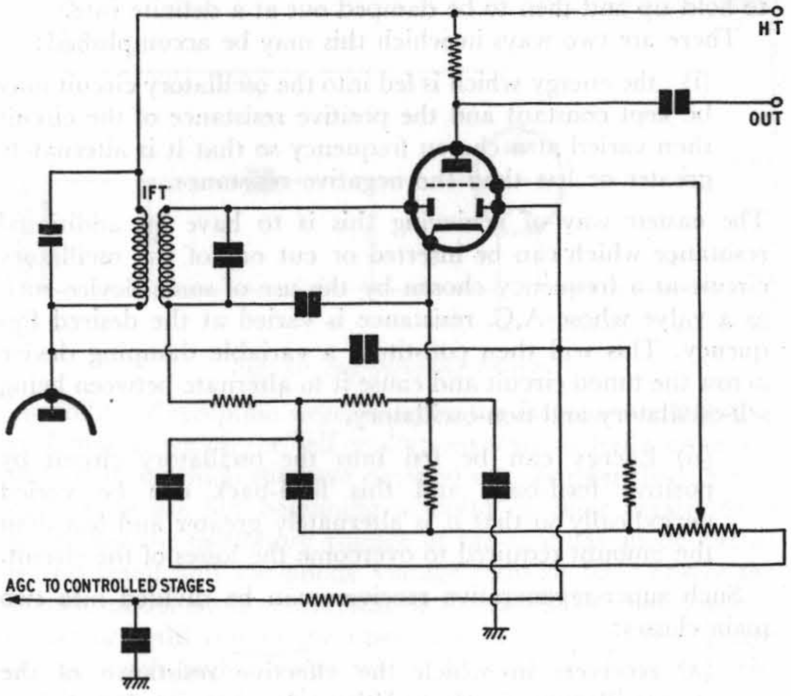


Fig. 55 AGC using double diode triode valve

of the signal will be necessary and the time constant will be short. Where the fading is slower then a large time constant will be advisable.

The time constant must not be too long or the AGC will not be able to follow the change in fading adequately.

Super-Regenerative Receiver

This receiver develops the principle of position feed-back and regeneration to its logical limit. A receiver with reaction control has already been mentioned. When reaction is applied to one stage the tuning is adjusted to a point where in the absence of an incoming signal the stage oscillates as near as possible to the frequency of the signal to be received. The effective resistance of the oscillatory circuit is varied deliberately at some specified frequency so that its value is alternately positive and negative. By this means free oscillations are periodically allowed to hold up and then to be damped out at a definite rate.

There are two ways in which this may be accomplished:

(i) the energy which is fed into the oscillatory circuit may be kept constant and the positive resistance of the circuit then varied at a chosen frequency so that it is alternately greater or less than the negative resistance.

The easiest way of achieving this is to have an additional resistance which can be inserted or cut out of the oscillatory circuit at a frequency chosen by the use of some device such as a valve whose A.C. resistance is varied at the desired frequency. This will then constitute a variable damping device across the tuned circuit and cause it to alternate between being self-oscillatory and non-oscillatory.

(ii) Energy can be fed into the oscillatory circuit by positive feed-back and this feed-back can be varied periodically so that it is alternately greater and less than the amount required to overcome the losses of the circuit.

Such super-regenerative receivers can be divided into two main classes:

(a) receivers in which the effective resistance of the oscillatory circuit and the valve associated with it is arranged to be self-adjusting so that self-oscillations alternately build up and are damped out without the need for a separate circuit to vary the reaction. These are called "self-quenching" receivers, or

(b) receivers where the periodic variation of effective resistance is brought about by means of a circuit other than that in which self-oscillations are alternately allowed to build up and die away.

These are known as "quench receivers".

Self-quenching Receivers

Self-oscillations are built up in this receiver and quenched periodically by automatically coupling the anode and the grid circuits. The circuit of a self-quenching receiver is similar to that of the reaction receiver and is shown in Fig. 56. The possibility of reception depends upon the fact that the negative grid bias at which the self-oscillations start to build up is less than that at which they are damped out, or quenched. When oscillations are just beginning to build up their amplitude is small, but the oscillatory voltage in the external anode circuit is also small and the anode voltage remains very nearly the same as under static conditions. As the oscillations build up the oscillatory grid voltage increases and the anode current jumps to the much higher value which is associated with the self-oscillations of large amplitude. If the negative grid bias is increased the anode current decreases and the oscillations

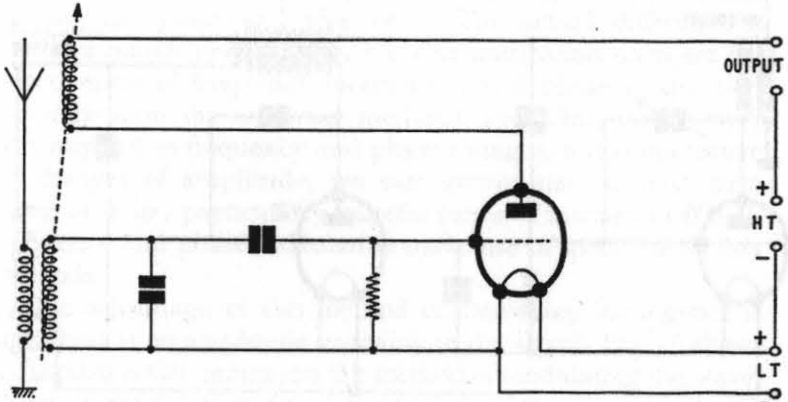


Fig. 56 Self-quenching receiver

decrease in amplitude. If the grid bias is made self-adjusting by using a condenser and leak a regular cycle of change can be induced. During the time that the condenser discharges the mean of grid potential becomes more and more negative and the oscillations die out. The whole cycle will be determined by the value of the condenser and the resistor.

In this receiver, although it exhibits extremely high gain, there is always a high background level of mush and noise.

Quench Receivers

Fig. 57 shows the circuit of a quench receiver. The signal is fed via the aerial coil into a tuned circuit. This circuit is tuned to the frequency of the incoming signal. A separate oscillatory circuit is maintained in self-oscillation at a high frequency; it is in fact a supersonic frequency of 20 kilocycles. This oscillator is coupled by the coil L and causes the anode voltage of V_1 to be modulated at the frequency of the oscillatory circuit of V_2. When the amplitude of the anode voltage rises free oscillations are built up; when it falls below the mean level the oscillations are damped, or quenched. When the signal frequency is received the gain at signal frequency is very large. These receivers have high background noise and suffer from distortion and instability. They have been used extensively in radar.

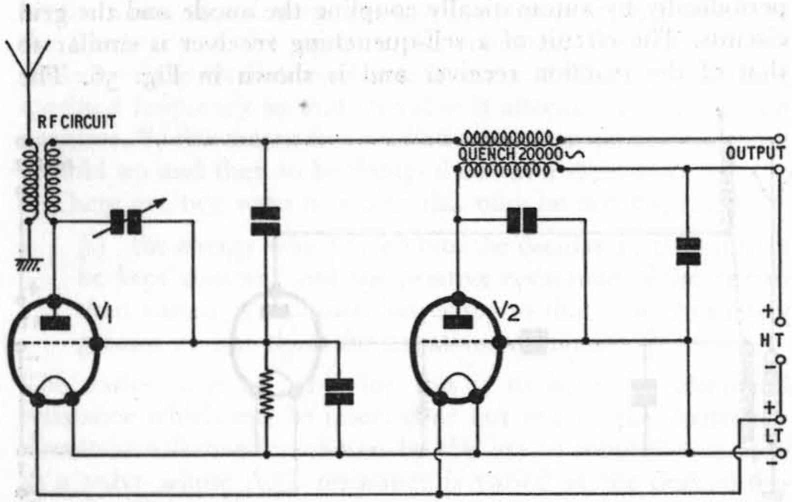


Fig. 57 Quench receiver

Frequency and Phase Modulation

So far the receivers described have all been concerned with amplitude modulation. However, it is possible to convey intelligence by modulating other properties of the carrier, including the frequency and the phase displacement. When the frequency of a carrier is varied in accordance with the variations of a modulating signal, the name "frequency modulation" is applied to it. Varying the phase of the carrier current is called "phase modulation".

Frequency modulation and phase modulation are not independent since the frequency cannot be varied without also varying the phase and vice versa. The actual difference is mostly a matter of definition. For communication purposes the effectiveness of frequency modulation and phase modulation depends upon the receiving methods. Providing the receiver will respond to frequency and phase changes, but is insensitive to changes of amplitude, we can discriminate against most forms of noise, particularly impulse forms of interference. Both frequency and phase modulation make use of special detection methods.

One advantage of this method of conveying intelligence is that there is no amplitude variation in the signal. Fig. 58 shows a diagram which represents the method of modulating the wave. When a modulating signal is applied to the carrier the frequency of the carrier is increased during one half-cycle of the modulating signal and decreased during the next half-cycle. The carrier will therefore vary between frequency limits alternately increasing and decreasing the frequency. The amount of change of the carrier frequency is called the "frequency deviation", and this is proportional to the instantaneous amplitude of the modulated signal. The deviation is therefore small when the amplitude of the modulating signal is small and great when the modulating signal reaches maximum amplitude, either positive or negative. The amplitude of the carrier as a whole does not increase or decrease during the process of modulation. If the phase of the current is changed in the circuit there is a frequency change during the time that the phase is being changed. The frequency deviation in this case depends upon how rapidly the phase shift is made.

In a properly designed system the amount of phase change is proportional to the amplitude of the modulating signal. The frequency deviation and phase modulation is proportional to both the amplitude and the frequency modulating signal. This is the difference between frequency modulation and phase modulation.

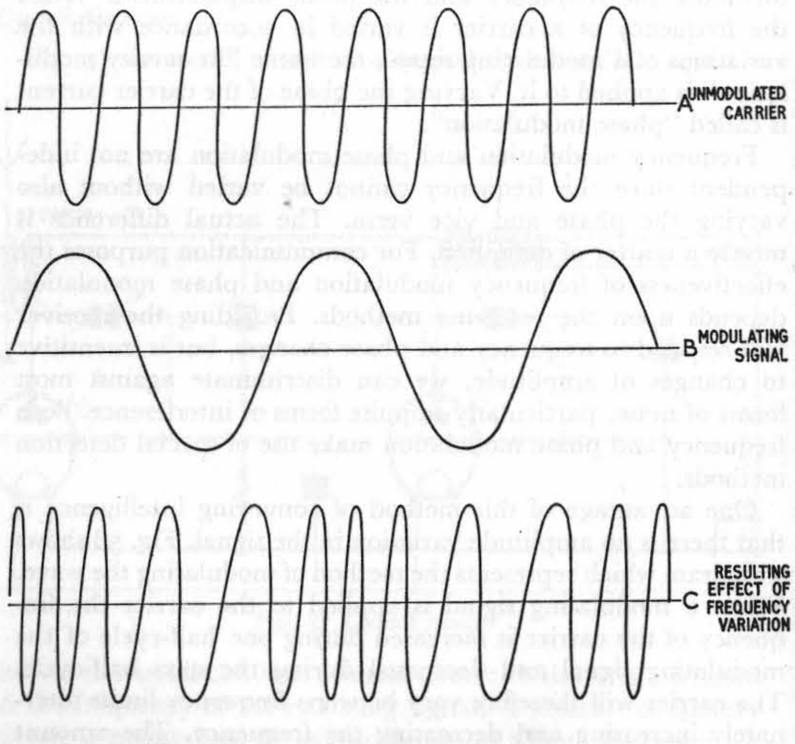


Fig. 58 Sequence illustrating frequency modulation

Frequency modulation and phase modulation are rarely used for radio astronomical purposes, so only this very brief mention of the alternative system is made.

Mention should be made of two other special types of receiver. These are not within the purse of the average amateur, but for the sake of completeness a brief description of them will be given.

The first of these is used at high frequencies and is called a "parametric amplifier". This is powered at radio frequencies and has a cyclic action driven by a section called a pump-oscillator. This amplifier is made to operate by successive pushes, as it were. The pump is a reservoir of energy which amplifies the original signal. A special crystal has been developed for this purpose and it can be used without low temperature techniques. There are a number of versions of this type of amplifier at work, and research into this type of amplifier has been extensive. Considerable improvement in noise figure is obtainable over the ordinary type of receiver and for high frequencies there is no doubt that this amplifier will come into wide use. The crystal, however, is extremely expensive as is also the type of parametric amplifier which operates with a travelling wave tube.

The other system makes use of low temperature techniques and is called the "Maser". The word maser was coined from a description of its mode of operation, which is, Micro-Amplification by Stimulated Emission of Radiation. One form of the maser uses a ruby crystal. This is kept at extremely low temperatures in a magnetic field. Under these conditions it is enabled to store energy. If left in a normal state it will give up this energy as radio frequency oscillations, but only slowly. When it is stimulated by an oscillator it will radiate very strongly. The energy is put into the crystal by a pump oscillator; the energy which is stored in the ruby is given up when it is stimulated by a radio wave. This will result in the energy being released many times greater than the original stimulated signal. This is a rather cumbersome and costly piece of apparatus and has to be mounted on the telescope itself. This leads to a certain amount of complication because the cooling method has to be that of liquid helium. It can only operate when the ruby is kept at this low temperature. For work on extremely high frequencies this receiver has, of course, done very valuable work. Its counterpart, the "Laser", makes use of this principle in dealing with light waves, and it may well be another method of exploring that part of the spectrum which lies between the infra-red, or the visible spectrum, and the beginning of our radio window of 1 centimetre.

Power Supplies

The object of a power supply is to provide means of obtaining pure direct current supply for the anode circuits and in certain special cases the grid circuits of receivers and other apparatus associated with the receiver. The heater of the valves used in the circuits can be fed with alternating current, though under special conditions it may be useful to supply these with direct current also.

The essential parts of a power supply are:

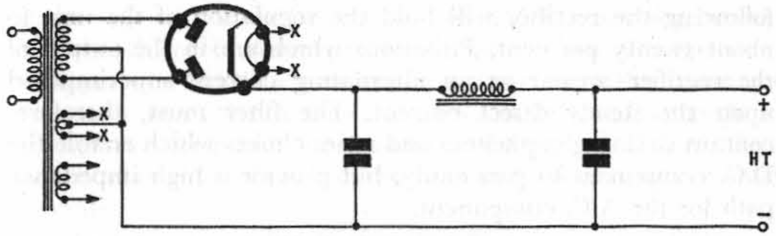
(i) a transformer which not only isolates the equipment from the power supply mains but also provides a means of obtaining different levels of voltage for rectification.

(ii) The rectifier may be a valve diode or it may be a metal rectifier such as copper-oxide or selenium.

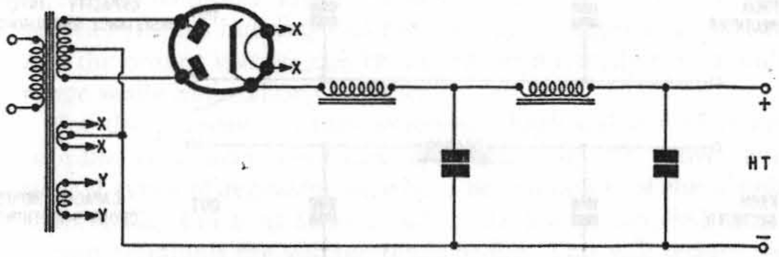
There are several modes of rectification: half-wave, full-wave and full-wave bridge rectifiers.

The principle of rectification with a diode has already been dealt with (pages 95–96). This was described as the half-wave rectifier. The shape of the rectified wave in this case has a considerable level of ripple, it is therefore more usual to use a full-wave rectification system. This will use two rectifiers, one of which will rectify the positive half-cycle of the alternating current and the other which will rectify the negative half-cycle of the current. Typical circuits are shown in Fig. 59*a*, *b*, *c*, *d*.

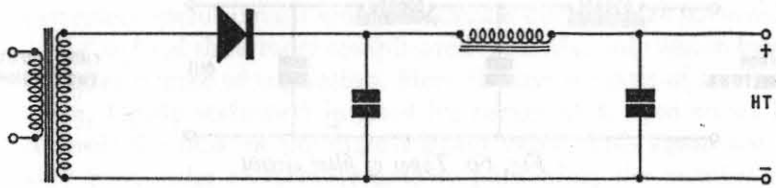
Two requirements of the power supply are, firstly, that the pulsating D.C. waves from the rectifiers shall be followed by filter circuits which substantially smooth out the pulsations. And secondly, that the voltage shall remain substantially constant over varying degrees of load. The output voltage regulation of the power supply tends to deteriorate as more current is drawn. This is because greater voltage drop occurs through the chokes and the rectifier during this time. Where an extremely high degree of regulation is a prime requirement, as in the case of radio astronomy, special arrangements in power supplies are made to hold the voltage within very narrow limits.



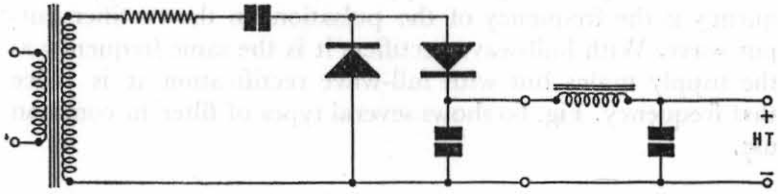
SIMPLE FULL WAVE RECTIFIER USING VALVE (A)



FULL WAVE WITH CHOKE INPUT FILTER (B)



HALF WAVE SYSTEM USING SELENIUM OR COPPER OXIDE RECTIFIER (C)



VOLTAGE DOUBLING USING TWO RECTIFIERS (D)

Fig. 59 Four examples of power rectifiers

Without special regulation systems a properly designed filter following the rectifier will hold the regulation of the unit to about twenty per cent. Pulsations which are in the output of the rectifier appear as an alternating current superimposed upon the steady direct current. The filter must, therefore, contain shunting capacitors and series chokes which enable the D.C. component to pass easily, but provide a high impedance path for the A.C. component.

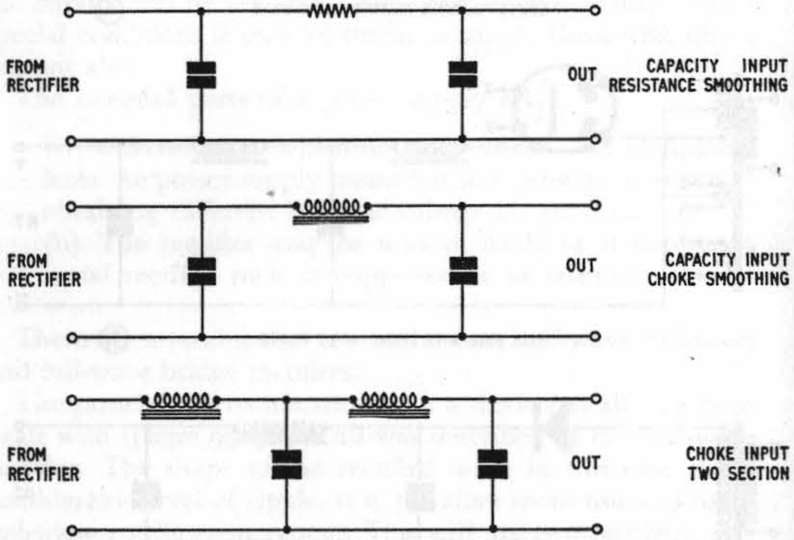


Fig. 60 Types of filter circuit

The alternating component of the power supply is called the "ripple" and the effectiveness of the filter can be described in terms of the percentage of ripple that is left. The ripple frequency is the frequency of the pulsations in the rectifier output wave. With half-wave rectifiers it is the same frequency as the supply mains but with full-wave rectification it is twice that frequency. Fig. 60 shows several types of filter in common use.

Voltage Stabilization

Gaseous regulator valves are useful to maintain voltage stabilization when an inexpensive method is required. The

regulator tubes consist essentially of a gas-filled valve which has a striking voltage or starting voltage. This starting voltage will be thirty to forty per cent higher than the operating voltage. The load is connected in parallel with this valve and for stable operation a minimum current of 5 milliamps to 10 milliamps is required for the valve. Voltage regulation of the order of one per cent can be obtained with such regulated circuits.

There are a number of circuits which have been developed for regulating the voltage output of a power supply using electronic methods. These are more complicated than the gas rectifier circuits but they can handle high voltages and currents and the output voltage can be varied continuously over a wide range while maintaining stability.

For the purpose of radio astronomy high stability of power supplies is a very necessary criterion. Fig. 61 shows four typical types of regulator supply. The operation of the simple supply, Fig. 61*a* is as follows: when the load is applied to the output terminals the voltage tends to fall. This will reduce the amount of grid bias on V allowing more current to flow and tending to restore the voltage to its original level. This is an extremely useful circuit where moderate currents are involved.

Fig. 61*b*,*c*,*d*, show more complicated types of supply which have a higher degree of regulation. Here the use is made of another valve, highly stabilized in itself by means of a neon tube. It controls the bias on the main current valve. This again has a very high order of stability and by paralleling the number of current valves at V the amount of current that the supply unit can deliver can be increased.

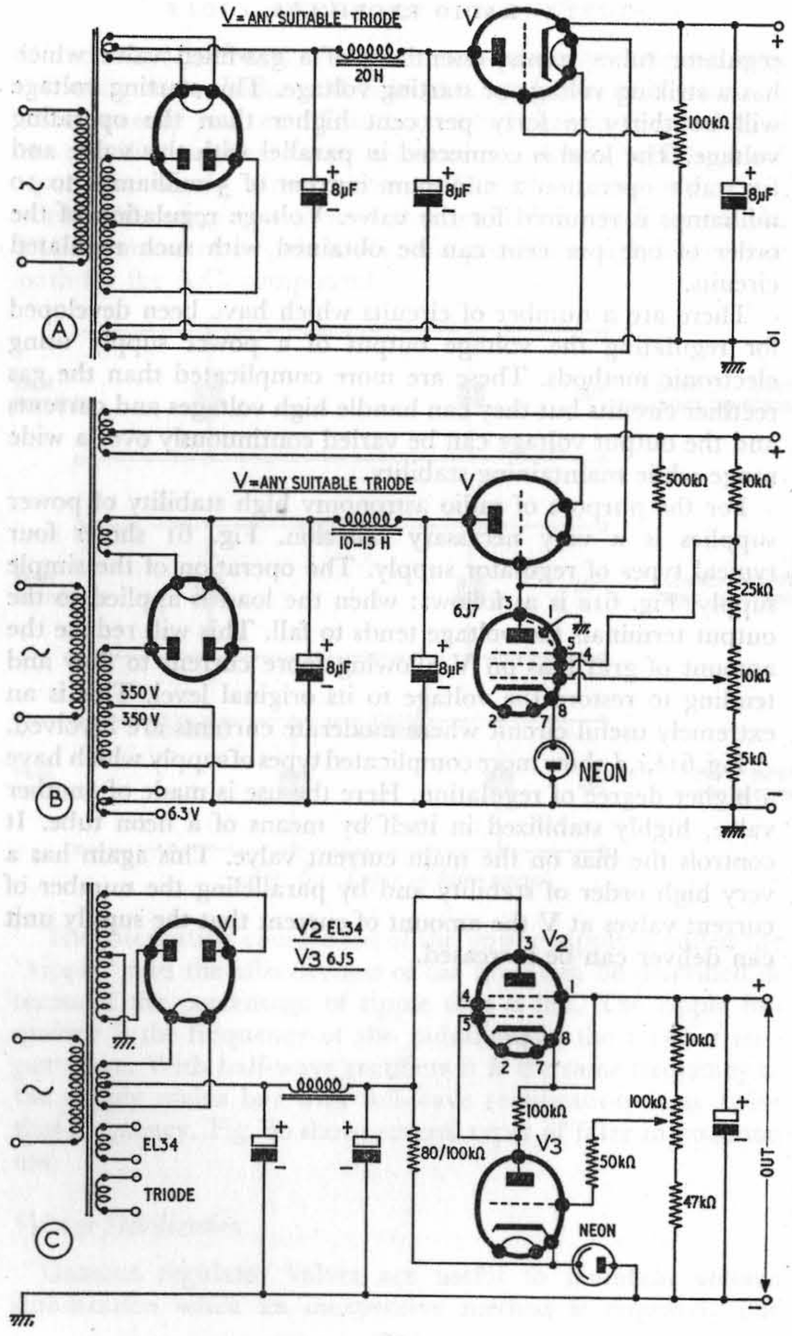


Fig. 61a, b, c Three regulated power supply circuits

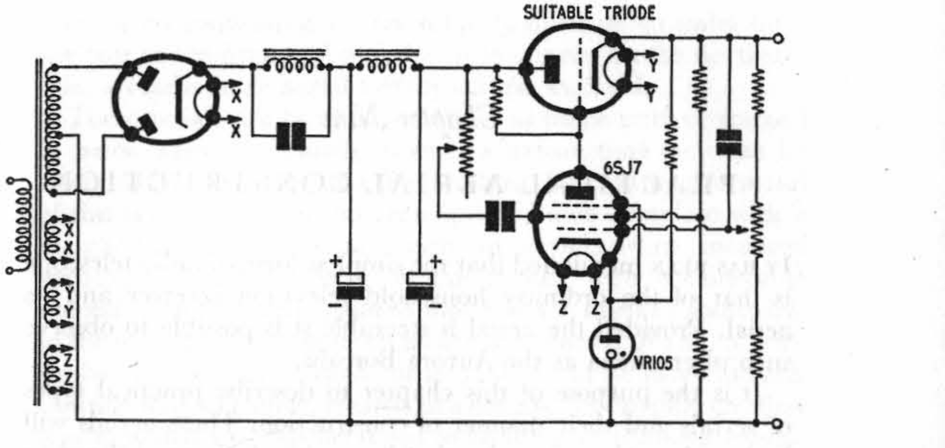


Fig. 61d A fourth method

Chapter Nine

PRACTICAL AERIAL CONSTRUCTION

IT HAS BEEN mentioned that the simplest form of radio telescope is that of the ordinary household television receiver and its aerial. Provided the aerial is steerable it is possible to observe such phenomena as the Aurora Borealis.

It is the purpose of this chapter to describe practical types of aerials and their manner of construction. These aerials will be based on the principles that have already been outlined in Chapter 7. Dimensional details of each aerial will be given and the instructions contained in this text together with the diagrams should enable the amateur to construct satisfactory systems for himself.

There are certain points to remember so far as the location is concerned. There must be a reasonably clear aspect to the south and preferably as far east and west as possible. If there are power supply lines obstructing the view and within 30 ft or 40 ft of the aerial system that is erected, allowance will have to be made for this obstruction; there may be spurious noises and other interference when such lines are within the beam of the aerial system. When working on the higher frequencies, trees and buildings, if within the distance quoted, will also cause some difficulty.

In the case of radiations from the Sun during its active periods, these will be sufficiently strong to break through local interference so that it is necessary to plot the local conditions, as a map can then be used to offset the results obtained when the aerial is pointed, for example, to the Sun.

The aerials will be described in the order in which they were taken in the earlier chapter. The first types to be dealt with will be the multi-element aerials.

There are certain commercial multi-element aerials for use in the television bands, particularly band three, which are suitable for radio astronomy. If a frequency is chosen which does not lie in the region of a local channel, and provided that

there is no transmitter nearer than 25 miles or 30 miles from the rear of the proposed site, that is to say not on the northerly side, a ready-made aerial system will be available.

These aerials can be obtained either as single units or stacked in pairs. They are usually of such a nature that they can be assembled in a matter of a few minutes. One particular aerial of this type is mounted on an elliptical boom complete with a matched coupling system. Such an aerial, when mounted vertically, has been used by the author with considerable success. The first arrangement made was a support somewhat in the manner of goal posts on a base constructed of 2 in. × 2 in. timber with wheels at each corner; this could be moved to any position and, of course, rotated in azimuth. A 2-in. pipe crossbar was provided and this arranged to be some 15 ft above ground level. Three groups of aerials were used, that is to say, three pairs mounted one wavelength apart horizontally on the crossbar. The crossbar could be rotated so that the aerials were oriented in altitude without difficulty.

The type of support described can be used for a number of different sorts of aerial. It was, therefore, standardized as a support. Reference to the diagram in Fig. 62 will show the general construction.

The base is made of 2 in. × 2 in. timbers bolted together. Wheels from 6-in. to 8-in. in diameter are attached at the four corners of the base. These can be pram wheels or scooter wheels, which are readily available at any handyman shop.

The uprights are also made of 2 in. × 2 in. material bolted into position with provision at the top for the fitting of a 2-in. diameter pipe for the support of the aerial. This pipe can very well be a standard television mast, since this offers a rigid member, but of reasonably light construction. Water pipe or gas pipe will be found rather heavy, although the standard scaffolding pole, such as used by building constructors, is quite suitable. This is usually of steel. Upon this may be mounted multi-element aerials of the Yagi type of multi-element aerial, with the reflector as in the Kooman array. The supporting assembly can also be used in the case of a corner reflector, providing the width between the uprights is made of the correct dimensions. If the aerials are chosen for a frequency in the

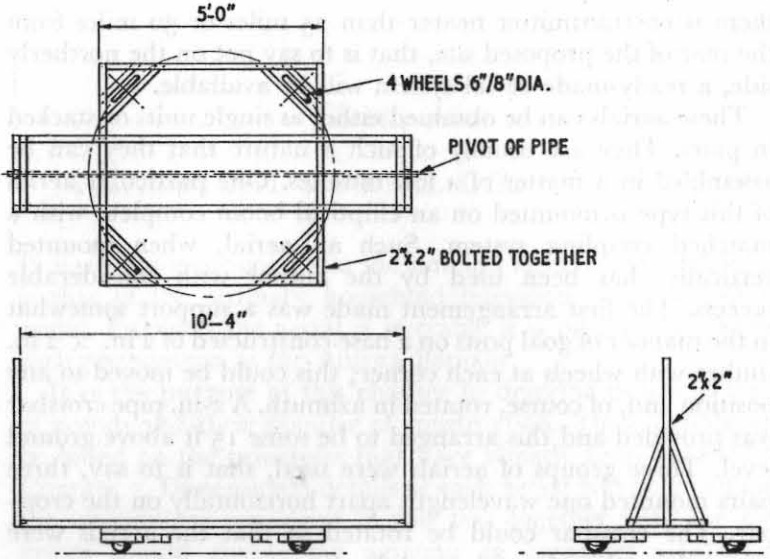


Fig. 62 Rotating base and supports for aerial assembly

region of band three channels and upward, this will still be a reasonably-sized assembly which can be used in a back garden. There is scope here, of course, for individual ingenuity.

The first steps to be taken before the construction of the aerials is attempted must be to decide the frequencies that are to be used for observation. Only one channel has been allocated to radio astronomy within the band available to the amateur and this is 137 megacycles. This is sufficiently near the amateur band of 144 megacycles to make similar equipment quite suitable for use in radio astronomy. At 108 megacycles we have satellite telemetry bands which differ from one another by small amounts. Reference to other works on radio astronomy will show that the frequencies around 80 megacycles have been used extensively, as have also 179 megacycles to 190 megacycles, 200 megacycles to 250 megacycles and 300 megacycles to 500 megacycles. For the purpose of this chapter, therefore, frequencies will be stated for each aerial and the dimensions given for the aerials will relate to those frequencies. In each type of aerial two frequencies will be dealt with and dimensions given

to suit. Where commercial aerials are used then the frequencies can lie anywhere in the five channels available in band three. For practical purposes in the average garden 80 megacycles is the lowest frequency used, but where a larger area of land can be used more ambitious schemes can be attempted. This would apply particularly in the case of the corner reflector to be described later, for if sufficient space were available then attempts could be made to record radiations from the direction of the planet Jupiter.

No instructions are included in this chapter for the construction of dish or paraboloid aerials, for though not beyond the ability of the amateur to construct either in steel or in wood, the size that can be achieved in the average garden is not sufficiently large for useful work. For example, a dish of 25 ft to 30 ft would be about the largest that could be accommodated in the back garden, provided always that permission could be obtained from the Local Authority. The resolution at frequencies available to the amateur, say up to 500 megacycles, would give at best a beamwidth of some 6 degrees. Since this is twelve times the diameter of the photosphere of the Sun, only extended sources could be examined, unless the aerial was combined with another as an interferometer. Such an elaborate arrangement would not be justified in the case of the amateur, for with the same amount of money he could build several of other types enabling him to cover a wider band of frequencies in a much more useful manner. It might be thought that a dish would be useful for the observation of satellites, but it would be a case of unnecessary refinement, as a simple Yagi aerial backed by a reflector or a helical aerial is all that is required to follow the progress of satellites for the normal work which would be undertaken by an amateur, that is doppler measurements.

The aerials to be described, therefore, are those within the reach of the amateur.

The most expensive equipment to be used would not exceed the cost of, say, a good camera and enlarger and the other associated equipment, or a complete ciné outfit. As a hobby, therefore, radio astronomy is no more expensive than most other hobbies which employ somewhat advanced techniques.

The first aerial that will be dealt with is a six-element multi-element aerial, or Yagi; this is set on a centre frequency of 80 megacycles. Reference to the diagram, Fig. 63*a*, will show the arrangement of the elements. The central boom or support can be a tube of aluminium or it can be a construction of timber, the very common form of timber available in the average supplier's yard is 2 in. × 1 in. rough sawn. This is extremely useful and, if selected to be reasonably clear of knots, will provide a basic element, as it were, for the amateur constructor. The suggested layout for the boom for the support of multi-elements is shown in Fig. 63*b*. No dimensions are given, for these will be arranged to suit the particular case.

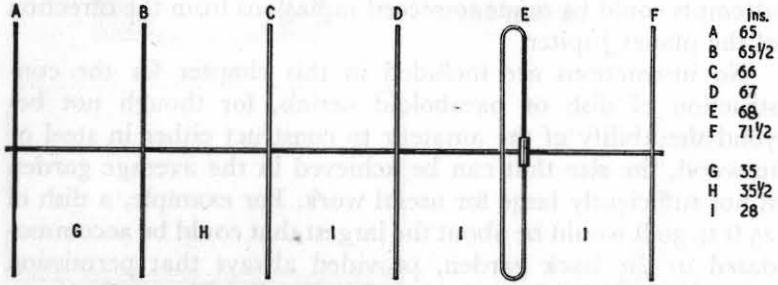


Fig. 63a Arrangement of elements

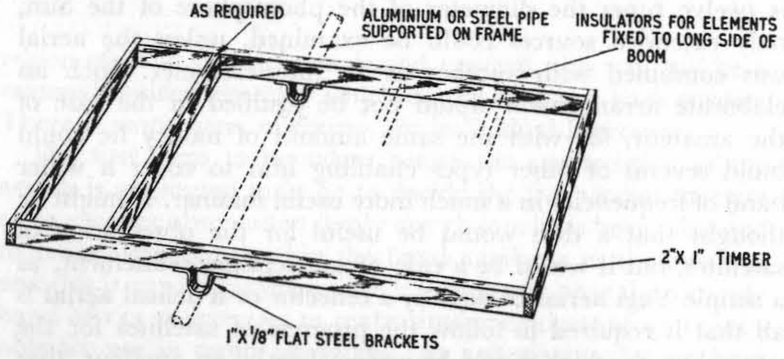


Fig. 63b Boom arrangement

The six-element aerial will be supported either on the frame construction which has been suggested, or alternatively it can be arranged on a mast which has an attachment to enable the aerial to be rotated in azimuth and also steered in altitude. The construction of the swivel joint can be fixed at the top of the mast and is shown in Fig. 63*c*. This is constructed of tube and slotted angle and will be found applicable either to the boom support made of timber or of aluminium tube.

The elements of the aerial can be of aluminium tube, copper tube or even stiff wire. It is recommended that aluminium tube of $\frac{3}{8}$-in. diameter should be used. If a folded dipole is used for the principal element, and this is recommended where possible, the dimensions would follow standard practice, that is to say that at this frequency the diameter of the rods would be $\frac{3}{8}$-in. to $\frac{1}{2}$-in. diameter and the spacing between the two sides of the dipole some $3\frac{1}{2}$ in. A detail of the dipole is shown in the diagram, Fig. 64.

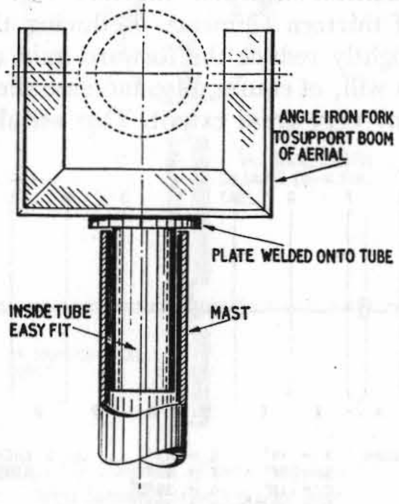


Fig. 63c Swivel

The design of this aerial is such that the central point of all the elements is neutral, so it can be bolted bodily to the tube without having any effect on the performance. The centre point of the continuous side of the dipole is also neutral and

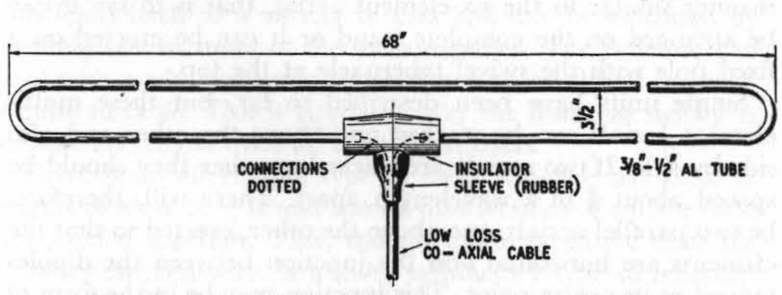


Fig. 64 Detail of folded dipole for six-element Yagi aerial

can be clamped direct to the boom. The alternative methods of matching have, of course, already been described – these are the gamma-match or delta-match.

The next multi-element aerial is designed to operate with a centre frequency of 137 megacycles. This is basically an eight-element aerial but it can be extended up to thirteen by the addition of further directors. Fig. 65 shows the complete aerial of thirteen elements. Reducing the number of directors will slightly reduce the forward gain and increase the bandwidth. It will, of course, also increase the beamwidth although not to any very great extent. This aerial can again be mounted in a manner similar to the six-element aerial, that is to say it can be arranged on the complete stand or it can be erected on a fixed pole with the swivel tabernacle at the top.

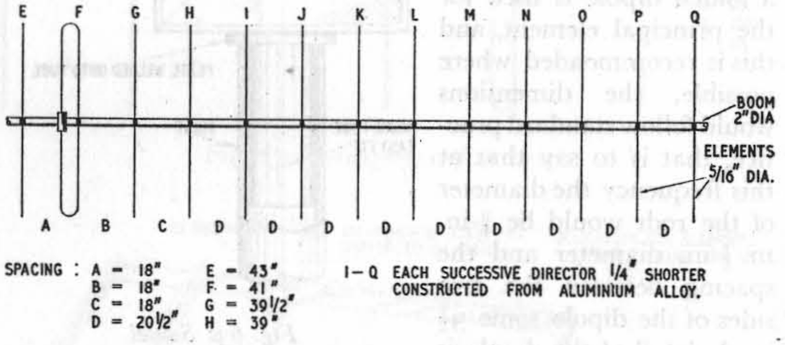


Fig. 65 Thirteen-element Yagi aerial

Single units have been described so far, but these multi-element aerials can be stacked one above the other and also side by side. If two aerials are stacked together they should be spaced about ⅜ of a wavelength apart. There will, therefore, be two parallel aerials, one above the other, erected so that the elements are horizontal and the junction between the dipoles tapped at its centre point. This junction may be in the form of stiff wire or it may be in the form of coaxial cable joined as shown in Fig. 66. In this case the length of the phasing lines will be an electrical half-wavelength, i.e. the velocity factor of the cable must be taken into account and the free-space wavelength multiplied by this factor when making the calculation. In the case of ordinary commercial 72 ohm coaxial cable this will be 0·65 for the solid di-electric cable and 0·84 for the air-spaced di-electric cable. When more than one pair of aerials are stacked side by side then the feeds from these matched pairs must again be matched and this will be done according to the methods described at the end of the chapter.

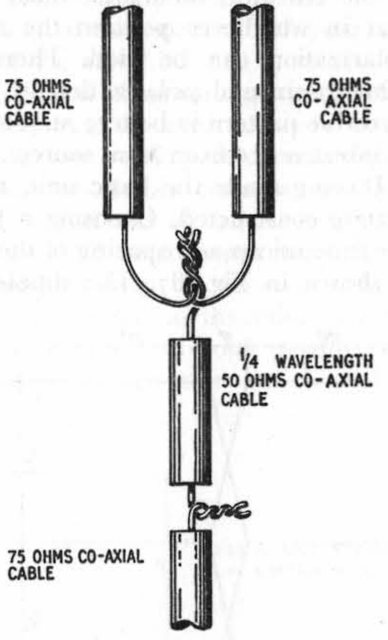


Fig. 66 Method of matching two 75 ohm inputs to one 75 ohm line. This can be multiplied when two or more groups are combined

Just as a standard base for most aerials is built up, so can a standard unit for all the other aerials be designed. The one to be described now makes use of a basic unit so chosen for its size that it can be accommodated in the average workshop, or at the worst carried into the kitchen. Its basic size is 10 ft × 6 ft. This lends itself to a variety of uses and can be extended by coupling one or more of the units together. The reflector can be of wire mesh, and ordinary chicken wire of 1½ to 2-in. mesh is quite suitable. This is stretched over the frame as tightly as possible keeping it even all over the surface.

There is an alternative type of mesh which is available in lengths of 25 ft × 3 ft and has its intersections of square form and welded together. This, though more expensive than the ordinary chicken wire, is a more satisfactory type of reflector. The mesh is fixed to the framework with ordinary wire staples. Thin wires can be used as reflectors and these should be spaced not more than one-eighth of a wavelength apart at the highest frequency to be used. A good standard spacing is 2-in.

If the full versatility of this unit is to be obtained using wires as the reflector, then these must be run in both directions so that in whichever position the reflector is placed horizontal polarization can be used. There are several advantages in using horizontal polarization in radio astronomy; one is the directive pattern is better; another is that it is less susceptible to interference from local sources.

Having made the basic unit, the first aerial can be immediately constructed. Choosing a frequency of 220 megacycles, the dimensions and spacing of the dipoles with the connections is shown in Fig. 67. The dipoles are supported on wooden brackets attached to the sides of the reflector. The distance between the dipoles and the reflecting screen should be between 0·2 of a wavelength and 0·25 of a wavelength. This may be used on its own since it may be erected on one edge with two supports at the back and the angle adjusted to face the source of extra-terrestrial radiations. In the case of the Sun it would be possible to set the position of this due south, for the beam-width is sufficiently wide for the Sun to be in the beam of the aerial for several hours as the aerial is carried round by the rotation of the Earth on its axis. During periods of active radiations from the Sun this simple arrangement will give quite a good account of itself.

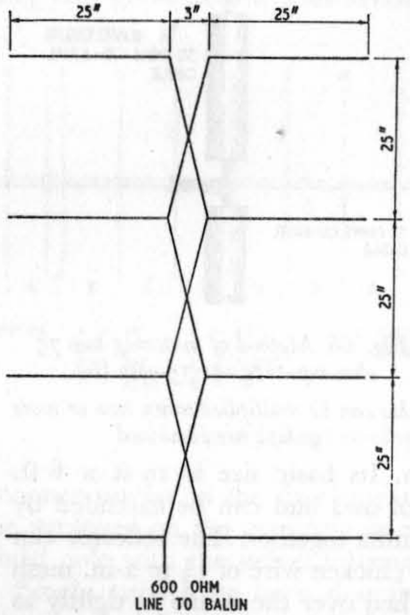
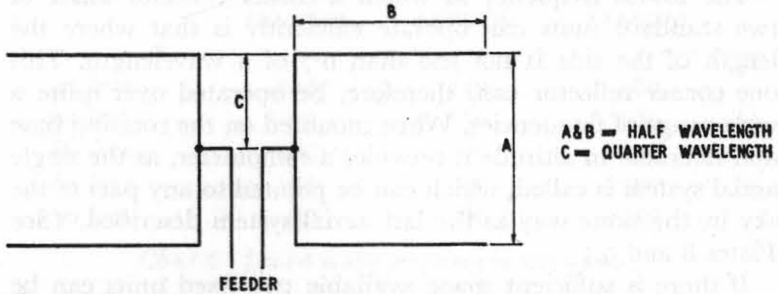


Fig. 67 Dimensions for dipoles for 220 megacycles

We can extend this idea by adding more banks of half-wave dipoles and erect them side by side. The beam will now be considerably narrowed and the Sun will remain in the beam for a shorter time. The gain of the aerial has, however, been very considerably increased. The vertical angle of the aerial need only be changed about four times a year since the beam-width in a vertical direction is sufficiently wide for the Sun to remain in it for a good portion of each part of the year, depending on the adjustment of the aerial.

If we use the alternative method of mounting and put the aerial on our standard rotating base we need only two of the units which, when bolted together, will provide a reflector 12 ft × 10 ft. At the frequency previously given we may use two sets of full-wave dipoles, or four sets of half-wave dipoles. To mount this on the base we shall need to put supports on the side of the reflector, and this may be done by adding lengths of 2 in. × 1 in. material bolted to the sides in a triangular formation to provide support. U-bolts or flat steel brackets can then be shaped to fit the pole and bolted or screwed to the main reflector assembly. This will be at the centre point where the two units are joined together. If four triangular stiffeners are arranged this will usually be found to make the whole assembly quite rigid, and this arrangement has the additional advantage that counterweights can be placed at the apex of the triangle and the whole system can be balanced for steering in altitude.



m/cs	A&B	C
80	35"	17"
85	30"	15"
90	28"	14"
95	25"	12½"

Chart 5 Dimensions of "lazy H" for 4 different frequencies

The same reflector can, of course, be used for lower frequencies in which case the number of aerials that can be accommodated is reduced. At 80 megacycles, for example, the type of aerial known as the "lazy-H" would be a suitable choice. The dimensions of the aerial elements themselves are shown in Chart 5 covering four different frequencies.

Corner Reflectors

We now turn to the corner reflector. Once again we make use of our standard unit and arrange this in the form of a right-angle. To each end we add cross-struts with another vertical strut which bisects the angle between the two sides of the reflector. This will serve two purposes:

(i) to support the ends of the reflector between the supports, and
(ii) to support the aerials which will be arranged horizontally at a specified distance from the apex of the assembly.

The lowest frequency at which a corner reflector made of two standard units can operate efficiently is that where the length of the side is not less than 0·7 of a wavelength. This one corner reflector can, therefore, be operated over quite a wide range of frequencies. When mounted on the rotating base and steerable in altitude it provides a radiometer, as the single aerial system is called, which can be pointed to any part of the sky in the same way as the last aerial system described. (See Plates 8 and 9.)

If there is sufficient space available two fixed units can be used on an east to west baseline and, provided the distance between them is ten wavelengths or more at the frequency being used, a useful interferometer can be set up. Since the aerials will be steerable in altitude it will be possible to scan strips of the radio sky. Successive 24-hour scans, with the Earth sweeping out a narrow strip the width of the beam of the aerial on each occasion, will enable a complete map to be built up.

If a frequency around 200 megacycles to 240 megacycles is chosen the whole interferometer can be accommodated in an average back garden, for only a distance of some 50 ft to 55 ft is involved between the extreme ends of the system. It is, of course, necessary to get as near to an east-west baseline as possible, although a deviation of up to 20 degrees will still allow useful work to be done.

The limiting range of frequencies is shown in Chart 6. The type of aerial used can be a single wire, or it can be in the form of folded dipoles. The choice will depend largely upon the number of aerials which can be arranged along the length of the reflector. Since the lowest frequency at which the standard unit can be operated is about 137 megacycles, that is the band allocated to radio astronomy, one full-wave dipole or two half-wave dipoles represent the choice available. The full-wave dipole will have a centre impedance of 600 ohms and it will be necessary to match this to the receiver. Two half-wave dipoles can be so arranged that matching can be done with coaxial cable. This will avoid the use of open wire feeders.

LENGTH OF BASELINE	5 WAVE-LENGTHS	10 WAVE-LENGTHS	20 WAVE-LENGTHS
40 feet	120	240	480
60 feet	75	150	300
80 feet	60	120	240
100 feet	50	100	200
150 feet	30	60	120

Chart 6 Lowest usable frequency in megacycles

The most satisfactory way to use the standard reflector is to choose a frequency where the ends of the aerials will be well inside the area of the reflector. As a recommendation, the frequency chosen should not be less than about 180 megacycles as this gives the optimum beamwidth in both vertical and horizontal directions. Particularly this should be the case where the aerial is to be used as a radiometer.

To accommodate a lower frequency it will be better to increase the number of units and arrange this so that the aerial can pivot on its apex, and so provide steering in altitude while being fixed in azimuth. If we take eight of our units and build up a large corner reflector with them each side will be 20 ft × 12 ft. We can now use this assembly at the lower frequency recommended of 80 megacycles.

The most important advantage of the corner reflector is that its front to back ratio is extremely high. Where the aerial is to be fixed and only occasionally moved in altitude it is perhaps the easiest of all aerials to assemble.

For frequencies other than those given in Chart 6 the following parameters should be observed when deciding on the frequency of operation.

The side of the reflector should not be less than 0·7 of a wavelength and the angle between them should be 90 degrees. The position of the aerial will be at the focal plane; the optimum position for this is at a distance of 0·35 wavelengths from the apex of the reflector. Suitable dimensions can then very easily be calculated for frequencies other than those that have been given in the text.

Example of large corner-type reflectors are shown in Plates 8 and 9. These aerials were designed primarily for a frequency of 27 megacycles, and are used for observations of the occultation of the Crab Nebula by the Sun's corona. This is the type of reflector also used for observing radiations from the planet Jupiter. The dimensions involved, however, are very large in relation to the average back garden since the height of the reflector at a frequency of 18 megacycles becomes 38 ft to 40 ft, and since the aerial must be at least two wavelengths long for satisfactory performance, its total length will be something over 80 ft.

The alternative for the amateur who wishes to attempt to make these observations is to use Yagi aerials which can be considerably reduced in size. However, it will still be quite a large assembly and even a six-element beam for this purpose is somewhat unwieldy since the dipole will be some 25 ft long and the assembly from back to front about the same size.

The alternative is to use a square loop with a reflector. This will reduce the overall size, since as each side of the loop is one-quarter wavelength these will be about 12 ft 6 in. However, it will still be several feet from back to front and will not operate very satisfactorily unless it is at least a half-wavelength above ground to the centre. The gain is no better than that of a three-element Yagi and unless two such aerials can be combined as an interferometer the back garden will tend to be a little small for this purpose.

There is no reason, however, why the square loop should not be tried on its own, providing the operation is carried out in

an area which is reasonably free from commercial interference. It is worthwhile to observe Jupiter at the time of transit on a frequency between 18 megacycles and 18·5 megacycles with a good commercial communications receiver.

The insulators that are used can be either the small egg-type insulator or, preferably, the 3-in. pyrex or ceramic insulator which are available in large numbers in the surplus market. Supports for aerials can very well be nylon cord since this weathers well and provides a non-conductive support. The use of chocolate blocks will facilitate connections to be made from one section to another.

Transmission Lines

An important part of the radio telescope is the method of connecting the aerial to the receiver. Here there are normally three methods from which to choose; coaxial cable, which is unbalanced line, and depending upon the type of cable will have a characteristic impedance of either 50 ohms to 52 ohms or 70 ohms to 80 ohms; parallel twin insulated cable with a characteristic impedance of 300 ohms; or open wire lines, which are usually spaced to provide a characteristic impedance of something of the order of 2,000 ohms, so when four are coupled together, as in the case of the Kooman arrays, each group is reduced to about 500 ohms and as further sections are paralleled so the impedance is reduced since they are connected in parallel.

The electrical length of the line will depend to a large extent upon its construction. This is always less than the wavelength in free space. It is necessary, therefore, to multiply the free-space wavelength dimension by a factor which is called the "velocity factor". Table 3 shows typical cables with the velocity factor that must be used when calculating sections.

With the coaxial line, standard fittings are available and are common with all television installations. A number of types are available from plain plug to the combination plug and connector. It is unlikely that the amateur will make much use of parallel lines for the reason that they are required to be mounted above ground level and sharp bends must be avoided, so a more or less direct route is necessary from the aerial to the

TABLE 3

TRANSMISSION LINES

TYPE OF LINE	IMPEDANCE OHMS	VELOCITY FACTOR V	ATTENUATION IN DB PER 100 FT FREQUENCY MEGACYCLES 21	28	50	144	220	420
Open wire	Varies with spacing	0·975	0·08	0·1	0·13	0·25	0·35	0·5
Coaxial air insulated	75	0·85	0·48	0·55	0·7	1·4	2·8	5·1
Coaxial cellular polythene	75	0·82	0·8	1·1	1·5	1·8	3·2	6
Coaxial solid di-electric single conductors	75	0·66	1·9	2·2	3·1	5·7	7·2	10·4
Coaxial solid di-electric stranded	75	0·66	0·98	1·15	1·55	2·8	3·4	4·9
Coaxial solid di-electric	52	0·66	0·83	0·98	1·35	2·5	3·3	4·8
Twin lead solid di-electric	75	0·82	3·6	4·1	5·5	9	11	15
Parallel conductors flat insulated	300	0·68	0·52	0·6	0·85	1·55	1·9	2·8

receiver. In coupling the aerial to the receiver some form of matching must be employed if the full energy received by the aerial is to be transferred to the receiver. Various forms of matching units can be made up, and where the aerial system is a balanced one it is always advisable to make some arrangement whereby the unbalanced line is properly coupled to the balanced aerial, since this will have a considerable effect on its performance.

The most usual method of coupling is by means of a unit called a "balun", and this is illustrated in Fig. 68. This enables an unbalanced coaxial line to be connected to a balanced line. When a number of coaxial cables are to be joined in parallel a matching system may be used so that if two incoming aerial feeders are to be matched together, as in the case of the simple drift interferometer, these are joined through a matching network using two types of cable. The length of the matching section will be one-quarter wavelength characteristic to that cable. The normal balun offers a transforming ratio of 4 to 1, therefore if the feed to the set is 75 ohms to 80 ohms coaxial this must be matched to an impedance of 300 ohms. In the case of the Kooman arrays, two stacks, each of 600 ohms joined together, will provide a terminal impedance of 300 ohms. A balun can now be connected directly to this. This will apply when full-wave dipoles are used. When half-wave dipoles are used, the characteristic impedance of the Kooman group will be about 75 ohms, so they will be coupled in the manner previously described for two 70 ohm to 80 ohm cables.

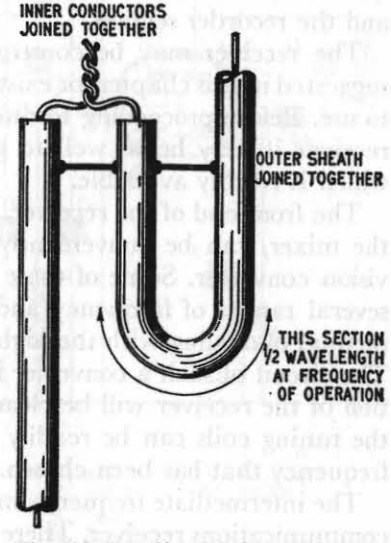


Fig. 68 Details of balun made up from co-axial cable

There are a number of methods of matching sections one to another. The simple forms which have been mentioned will be sufficient for most amateur purposes.

Chapter Ten

RECEIVER CONSTRUCTION

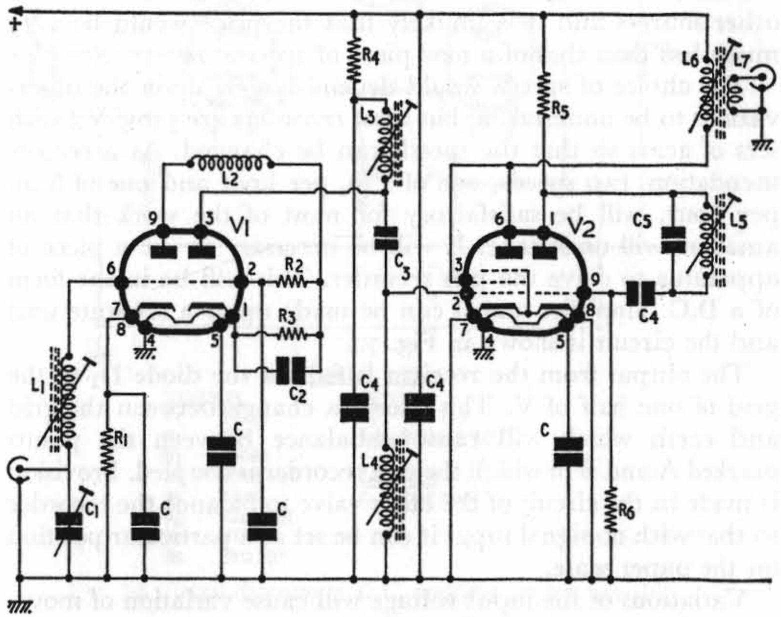
THE CHOICE OF a receiver will depend upon the frequency chosen for the telescope. The sequence will be the radio frequency and mixer section, the intermediate amplifier section, and the recorder section.

The receiver may be constructed according to the circuits suggested in this chapter, or existing apparatus can be converted to use. Before proceeding to discuss the design of the complete receiver it may be as well to go over the existing apparatus which is readily available.

The front end of the receiver, that is the radio frequency and the mixer, can be conveniently provided by a standard television converter. Some of these are available to accommodate several ranges of frequency and such a unit provides an easy method of dealing with the high frequency side of the receiver. The circuit of such a converter is shown in Fig. 69. The operation of the receiver will be clear from the circuit diagram and the tuning coils can be readily adjusted to suit the particular frequency that has been chosen.

The intermediate frequency may very well be a good standard communications receiver. There are a number of these available such as the CR100, the AR88, the R208, the S27, the CR150 and the national HRO. Each of these communications receivers will have its own audio output, so it is possible to use this to record directly on tape the output noise created by the voltages received at the aerial. This is a simple method which can be put into immediate operation since the tape recorder can then be played back and analysed at leisure. Thus it will be possible for the recorder to be switched on at the time that observation is required and allowed to run to the extent of the tape. This can be arranged either by a time-clock, or by a helper who switches on at the required time. It is not necessary for the observer to be present at the time of observation.

One advantage of using the standard communication receiver as the intermediate frequency and output section of the receiver, is that these receivers are well designed, very stable and reliable. Accurate tuning is available and usually there are certain other facilities which can be useful where it is required to narrow the band of operation. The ideal method of recording is, of course, with a pen recorder as permanent records can then be made on the paper roll and a receiver equipped with a pen recorder can be allowed to run throughout the twenty-four hours without attention. It is only necessary to read off from the chart whatever activity has taken place during that period.



C	·001pF Feed through	R_1	47Ω	V_1	ECC84
C_1	30pF Trimmer	R_2	18kΩ	V_2	ECF80
C_2	100pF	R_3	190Ω		
C_3	3pF	R_4	820Ω		
C_4	10pF	R_5	8·2kΩ		
C_5	5·6pF	R_6	10kΩ		

$L_1L_3L_4$	Coils resonate at 200/220m/cs
L_2	Cascode interstage coil
L_5	Oscillator coil
L_6L_7	To suit I.F. frequency chosen

Fig. 69 Converter for use with communications receiver

The pen recorder is essentially a meter which has had its movement specially designed to carry a pen which will mark upon the paper roll. A wide range of pen recorders are available and it is recommended that they are purchased direct from the manufacturer. It is not often that they are available from other sources and it is unlikely that the price would be very much less than that of a new piece of apparatus.

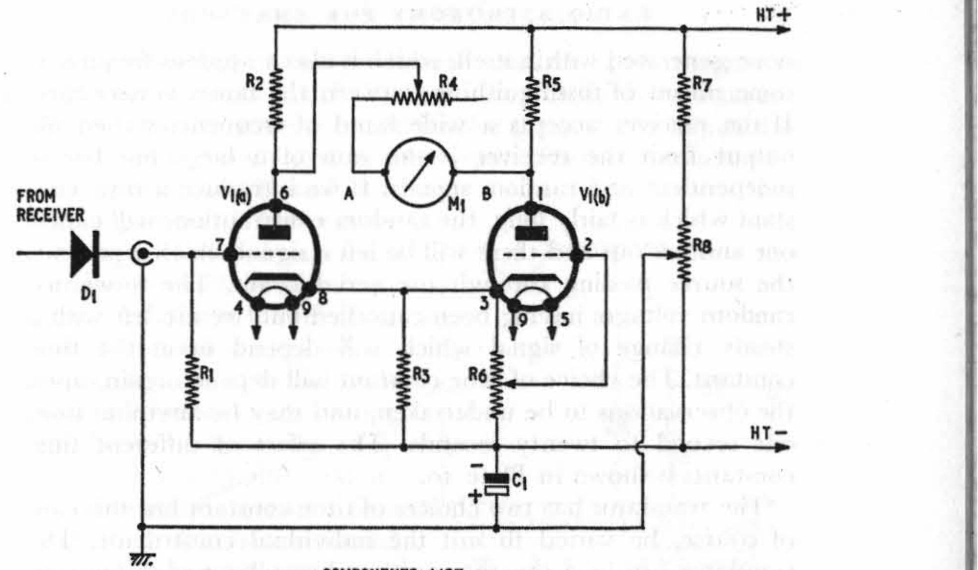
The choice of speeds would depend largely upon the observations to be undertaken, but most recorders are provided with sets of gears so that the speeds can be changed. As a recommendation, two speeds, one of 1 in. per hour and one of 6 in. per hour, will be satisfactory for most of the work that an amateur will undertake. It will be necessary to use a piece of apparatus to drive the pen recorder. This will be in the form of a D.C. amplifier and it can be made up as a separate unit and the circuit is shown in Fig. 70.

The output from the receiver is fed via the diode D_1 to the grid of one half of V. This causes a change between the grid and earth which will cause unbalance between the points marked A and B to which the pen recorder is coupled. Provision is made in the circuit of the other valve to balance the recorder so that with no signal input it can be set at a particular position on the paper scale.

Variations of the input voltage will cause variation of movement of the pen. When the signal rises due to a source passing through the beam of the aerial a rise will appear in the record. Since the speed of the paper is known, the time of this occurrence can be determined.

An alternative piece of apparatus which is simple and inexpensive to construct was designed by the author and is called a "translator". This consists of a diode detection system followed by amplification of the rectified voltage. Referring to the circuit in Fig. 71 the input is via the control R_3 and this controls the amount of signal applied to the diode. The rectified signal is applied to the grid of the following valve into the cathode circuit of which is connected the pen recorder. Variations of the input will therefore cause variations to appear on the chart of the recorder.

Now, the radiations which are being received by the aerial are at random frequencies, and since the receiver itself has



COMPONENTS LIST

RESISTORS		METER	
R_1	20 MΩ	M_1	1mA f.s.d
R_2	5kΩ	V_1(a), (b)	12AX7/ECC83
R_3	500kΩ		
R_4	20kΩ POT	**CAPACITOR**	
R_5	5KΩ	C_1	2μF 50wV
R_6	25kΩ POT		
R_7	47kΩ		
R_8	5kΩ POT		

Fig. 70 Circuit diagram of D.C. amplifier for pen recorder

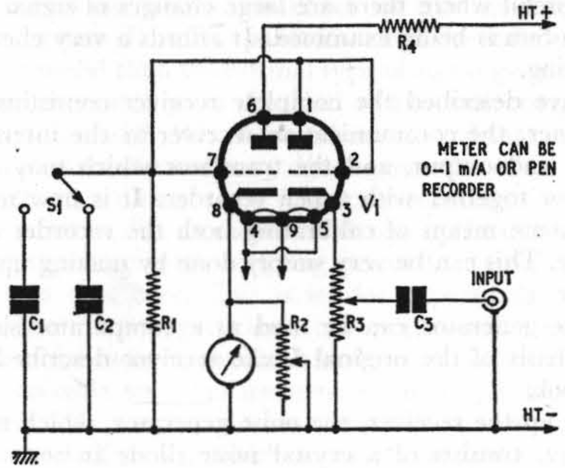


Fig. 71 Circuit of the translator unit

noise generated within itself, which is also a random frequency, some means of distinguishing between the noises is necessary. If the receiver accepts a wide band of frequencies then the output from the receiver is the sum of a large number of independent and random signals. If we introduce a time constant which is fairly long, the random contributions will cancel one another out and there will be left a signal which represents the source passing through the aerial beam. The unwanted random voltages having been cancelled out, we are left with a steady change of signal which will depend upon the time constant. The choice of time constant will depend, again, upon the observations to be undertaken, and may be anything from one second to twenty seconds. The effect of different time constants is shown in Plate 10.

The translator has two choices of time constant but this can, of course, be varied to suit the individual constructor. The translator can be a separate unit, and may be used to measure directly the output from the receiver when no pen recorder is available. It can also be used to examine the recordings which have been made on the tape recorder. In this case the pen recorder is placed by an ordinary meter. The variations of the position of the pointer can then be plotted at intervals of say five seconds. This will enable a graph to be built up of the change of level of signal. Though rather laborious it can still be very useful where there are large changes of signal such as when the Sun is being examined. It affords a very cheap form of recording.

We have described the complete receiver consisting of the turret tuner, the communication receiver as the intermediate amplifier and output, and the translator which may be used by itself or together with a pen recorder. It is now necessary to have some means of calibrating both the recorder and the translator. This can be very simply done by making up a noise generator.

A noise generator can be used as a comparator since this was the basis of the original Dicke receiver described earlier in the book.

To set up the receiver, the noise generator, which is shown in Fig. 72, consists of a crystal noise diode in series with a control and a meter for setting the level. The receiver is tuned

to the frequency that is to be used and the pen recorder or other meter set to a definite level. The aerial is now disconnected and the output from the noise generator is adjusted so that there is a slight rise in the output. The noise generator is now disconnected and the reading observed. The receiver output is increased to bring it to the level set by the noise generator; this is the reference point. Any deviation of increase or decrease will now bear a relation to the setting on the noise generator. Frequent checks in this manner will enable the observer to note how much drift is contributed by the receiver and how much from extra-terrestrial sources. It is generally advisable to check the receiver at the beginning and at the end of the period of observation, if this is short. If long-term observations are being undertaken then calibration once a day would normally be sufficient.

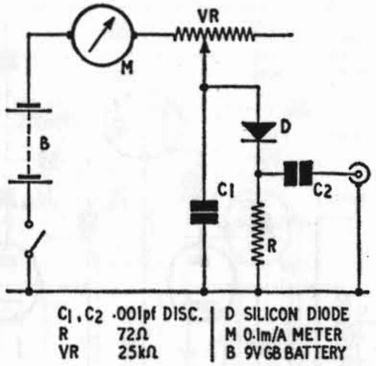
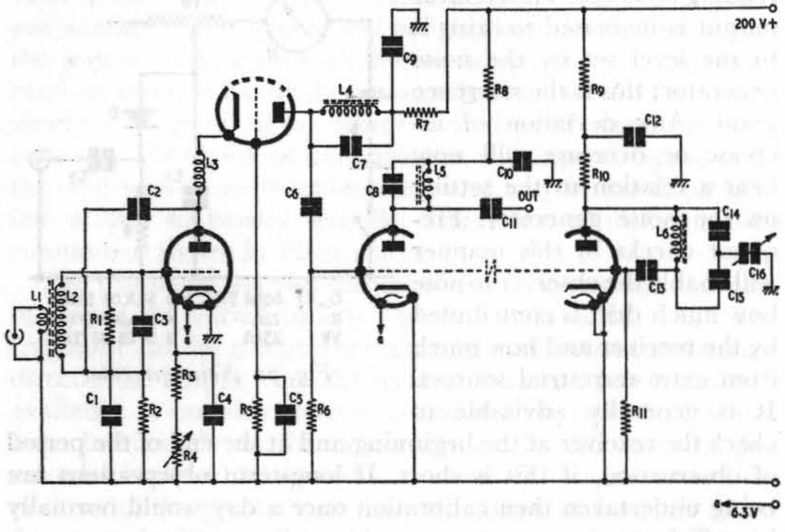


Fig. 72 Noise generator

A noise generator is an extremely useful piece of apparatus and provides an immediate check of the condition of the receiving system. From the point of view of the radio astronomer it is more useful than the normal type of signal generator, since the noise generator gives the kind of random voltages which we encounter with the radio telescope and has thus great advantages over the ordinary signal generator which is modulated with an ordinary radio frequency source.

The circuit in Fig. 73 shows the design of the complete radio receiver; input is via the coupling circuit L_1, L_2 which is tuned by an iron dust core. This is set for a particular frequency coverage from 210 megacycles to 220 megacycles. A slight modification of the radio frequency section is all that is required to make it suitable for other frequencies for which the aerial has been designed. The input from the tuned circuit is fed to the grid of one half of V_{1a} which is cascode-coupled to the other half V_{1b}. This valve is a low noise PCC 89. From the output of the second half of the valve the signal is fed via C_7 to the grid of the mixer V_2. The oscillator is half of the double triode which is an ECC 81. The coupling between the oscillator and the mixer is electronic and is of sufficiently loose coupling to avoid pulling by the oscillator. The intermediate frequency chosen is 38 megacycles. The output from the mixer is fed to the grid of V_1 and this is amplified by V_1 and passed on to four other stages, V_2, V_3, V_4, and V_5. The output from V_5 is by the transformer to provide a low impedance output and this is fed directly to the diode detector which is part of the pen recorder unit already described.



L_1 1¾T	22swg	R_1 4·7kΩ	R_7 0·1kΩ	C_1 33pf	C_9 1000pf
L_2 3½	PVC	R_2 47kΩ	R_8 4·7kΩ	C_2 ·2pf	C_{10} 001µF
L_3 4T	ins.	R_3 47Ω	R_9 3·3kΩ	C_3 2pf	C_{11} 47pf
L_4 1T	14swg	R_4 5kΩ Vari	R_{10} 3·3kΩ	C_4 0·001µF	C_{12} ·002pf
*L_5 9T	26swg	R_5 150kΩ	R_{11} 10kΩ	C_5 0·001µF	C_{13} 10pf
L_6 2T	22swg	R_6 100kΩ		C_6 5pf	C_{14} 10pf
	PVC ins.			C_7 4·7pf	C_{15} 5pf
				C_8 15pf	C_{16} 3–30pf

*For 16mcs. I.F. or aerial input of communications rec. the number of turns must be increased to 17 and adjusted.

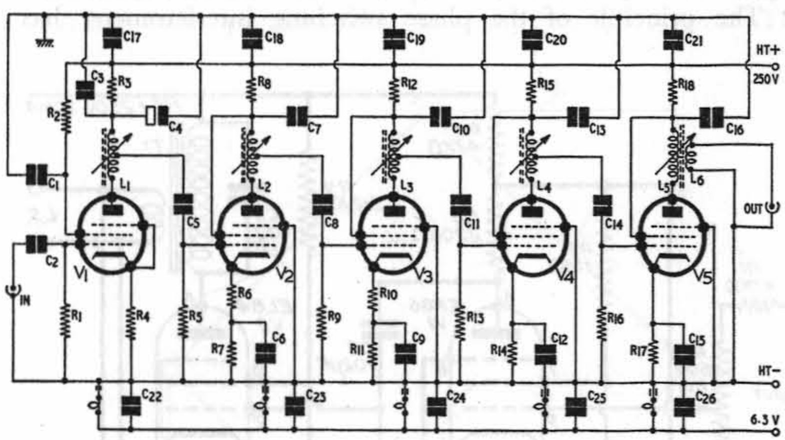
V1A & B	PCC89 or V2521
V2A & B	ECC81

R.F. Unit

Fig. 73a Complete receiver. R.F. unit

The tapping is provided for further audio amplification which is for the purpose of monitoring and this is the simple circuit shown in Fig. 74.

The layout of the radio frequency amplifier is important and each stage should be well screened, one from the other. Careful routing of wiring will ensure adequate gain with stability.



C_1	·001µF	C_9	·002µF	C_{17}	·002µF	C_{25}	·002µF
C_2	47pf	C_{10}	·003µF	C_{18}		C_{26}	·002µF
C_3	·002µF	C_{11}	·001µF	C_{19}		L_1	9T.TAP3
C_4	8µF EL	C_{12}	·002µF	C_{20}		L_2	9T.TAP3
C_5	·001µF	C_{13}	·003µF	C_{21}		L_3	9T.TAP3
C_6	·002µF	C_{14}	·001µF	C_{22}	·001µF	L_4	9T.TAP3
C_7	·002µF	C_{15}	·002µF	C_{23}	·002µF	L_5	12T.
C_8	·001µF	C_{16}	·003µF	C_{24}	·002µF	L_6	7T.

R_1	47kΩ	R_9	22kΩ	R_{17}	150Ω
R_2	33kΩ	R_{10}	33Ω	R_{18}	470Ω
R_3	2·2kΩ	R_{11}	68Ω		
R_4	180Ω	R_{12}	2·2Ω	V_1	6AK5
R_5	22kΩ	R_{13}	22kΩ	V_2	6BW7
R_6	33Ω	R_{14}	100Ω	V_3	
R_7	68Ω	R_{15}	2·2kΩ	V_4	
R_8	2·2kΩ	R_{16}	22kΩ	V_5	

Fig. 73b Complete receiver. Intermediate radio frequency section

A number of power supply units have already been described, but of course a commercial type of power supply could very well form the basis of a stabilized unit. Such power supplies are available quite cheaply and are extremely rugged. The addition of stabilization is a comparatively simple matter for the constructor and it is suggested that one or other of the designs which have already been given be adopted.

For those who wish to attempt interferometer work the extra apparatus required for phase switching will now be described.

The principle of the phase switching interferometer has

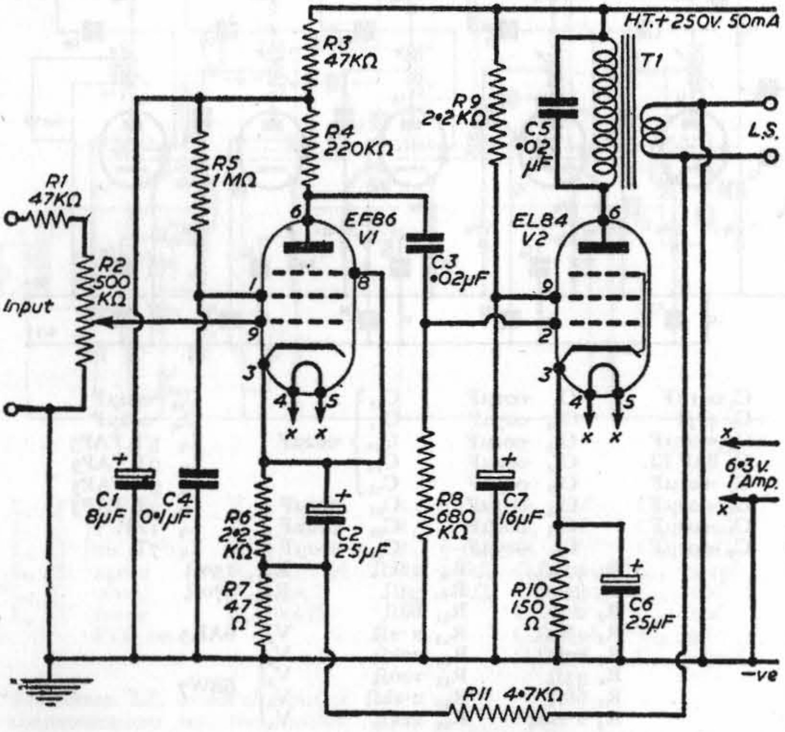


Fig. 74 Simple audio frequency amplifier

already been made clear. The extra units which are required are the phase switch, switch frequency generator and its driver, and a phase sensitive detector. The phase switch is the unit which inserts and removes the extra half-wave line. This is accomplished by means of a hybrid bridge, and is constructed of coaxial cable. The lengths of cable involved will again depend upon the frequency at which the telescope is to operate.

Reference to the diagram in Fig. 75 will show that three sides of the bridge are formed of coaxial cable a quarter-wavelength (electrical wavelength) in length, the fourth side is increased by a half-wavelength making three-quarters of a wavelength total. On the input there is a quarter-wavelength stub which is earthed and provides the necessary transformation. The

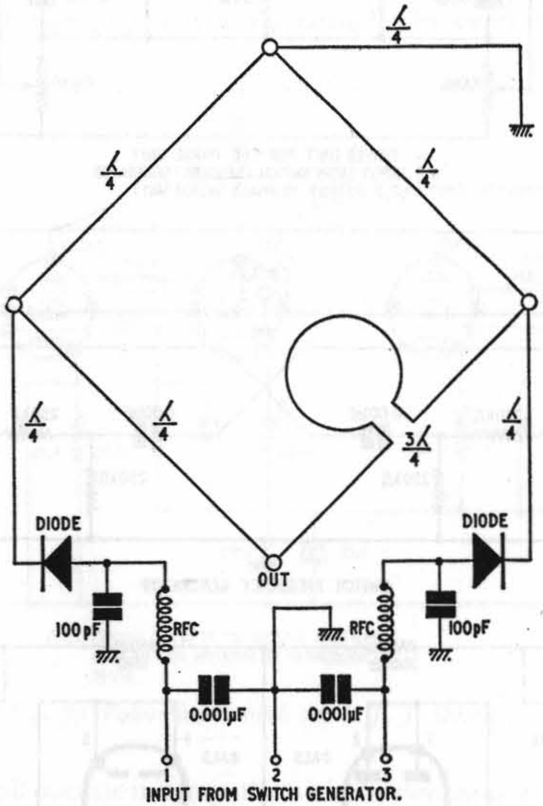
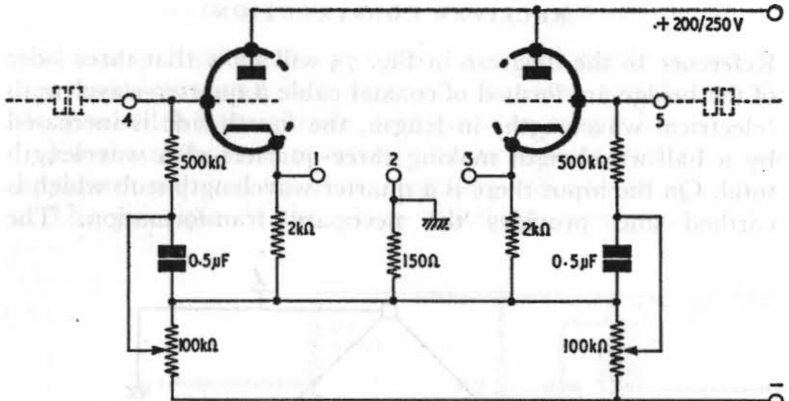
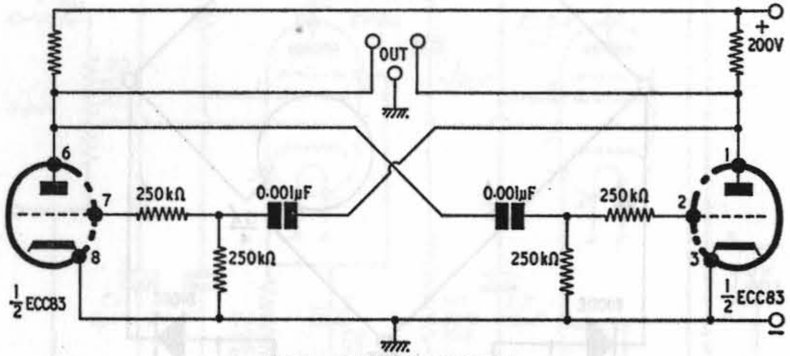


Fig. 75 Details of phase switch. Arms and stubs are made of co-axial cable

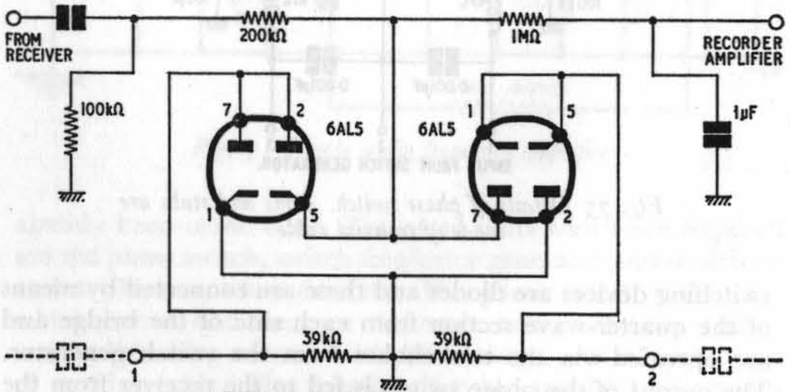
switching devices are diodes and these are connected by means of the quarter-wave section from each side of the bridge and are operated via the two chokes from the switch generator. The output of the phase switch is fed to the receiver from the point marked "out". When the switch is in operation driven



DRIVER UNIT FOR THE PHASE UNIT
4,5 INPUT FROM SWITCH FREQUENCY GENERATOR
1,2,3 OUTPUT TO PHASE SWITCH UNIT



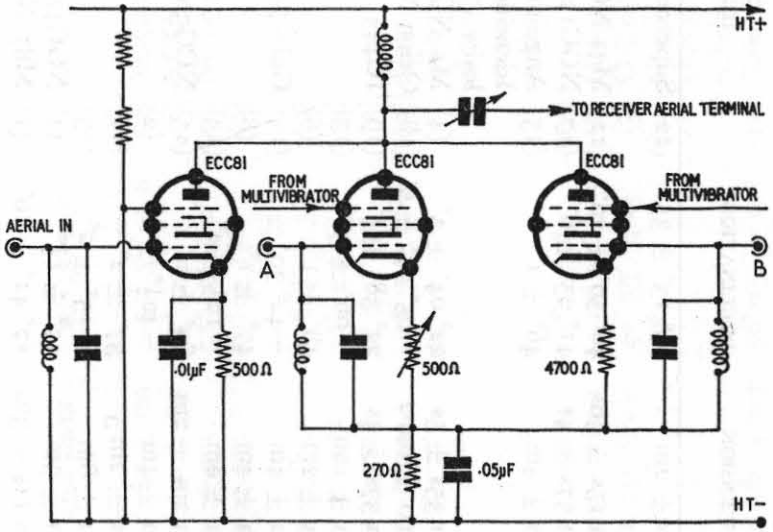
SWITCH FREQUENCY GENERATOR



CIRCUIT OF THE PHASE SENSITIVE DETECTOR
1,2 FROM DRIVE UNIT

Fig. 76 Details of units for phase switching

by the switch generator, this alternately closes one section or the other as the diode is blocked off. When diode 1 is blocked off the signal from the aerial to which the phase switch is connected is via the quarter-wave section and the three-quarter wave section. When diode 2 is closed by the switch generator the path of the signal is via the two quarter-wave sections. Thus with every change of the switch generator the path of the signal is changed by half a wavelength. In order that the



A.B. COAXIAL SOCKETS TO RECEIVE HALFWAY PHASING LINK APPROPRIATE TO FREQUENCY IN USE

Fig. 77 Valve phase switch due to J. J. Davies

receiver shall operate in sympathy with the switching generator, part of the generator is fed to the phase sensitive detector. This is switched on and off alternately in sympathy with the phase switch. The phase sensitive detector consists of two double diodes connected back to back and the phase sensitive detector is connected between the output of the receiver and the recorder amplifier. The switch frequency generator is a multi-vibrator; it is maintained in oscillation by feed-back from anode to grid. The frequency of the multi-vibrator can be chosen to suit the

LIST OF RADIO SOURCES

TABLE 4

CONSTELLATION	I.A.U. NO.	RIGHT ASCENSION	DECLINATION	IDENTIFICATION
Cassiopeia	00N6A	00h 22m ± 2m	64° 15′ ± 35′	(12) Supernova 1572 (no visible remnants)
Andromeda	00N4A	00h 40m 15s ± 30s	40° 50′ ± 20′	(11) M31 NGC224
Perseus	03N4A	03h 16m 37s ± 4s	41° 25′ ± 6′	(17) NGC1275 Colliding galaxies
Auriga	04N4A	05h 08m ± 4m	46° ± 1°	(13) Angular size of source is 1°4′ Identified with galactic nebulosity
Taurus	05N2A	05h 31m 35s ± 5s	22° 04′ ± 5′	(4) M1 NGC1952 Crab Nebula
Orion	05S0A	05h 33m·0 ± 0m·2	−05° 27′ ± 5′	(18) Orion Nebula M42 NGC1976
Gemini	06N2A	06h 13m 37s ± 4s	22° 38′ ± 5′	(19) IC443
Monoceros	08S0A	08h 08m ± 10m	−06° ± 30°	(13)
Lynx	08N4A	08h 09m ± 2m	48° ± 1°	(8)
Puppis	08S4A	08h 20m ± 4m	−42° 30′ ± 1°	(13) Galactic nebulosity
Lynx	09N4A	09h 16m ± 4m	47° ± 1°	(8)
Hydra	09S1A	09h 16m ± 2m	−12° ± 2°	(13)
Ursa Major	09N6A	09h 51m 20s ± 2m	69° ± 1°	(15) NCG3031
Vela	10S4A	10h 10m ± 4m	−42½° ± 20′	(2)
Ursa Major	10N5B	10h 30m ± 2m·5	57° ± 2°	(11)
Crater	11S1A	11h 38m ± 8m	−15° ± 2°	(13)
Canes Venatici	12N4A	12h 15m ± 3m	47° ± 1½°	(11) NGC4258
Virgo	12N1A	12h 28m 11s ± 37s	12° 41′ ± 10′	(1) M87 NGC4486
Centaurus	13S4A	12h 22m 24s ± 1m	−42° 37′ ± 8′	(1) NGC5128
Canes Venatici	13N4A	13h 27m 30s ± 3m	47° ± 1°	(11) NGC5195
Boötes	14N5A	14h 10m ± 2m	51° 30′ ± 1°	(11) NGC5457
Serpens Caput	15N1A	15h 10m ± 4m	11° ± 1½°	(2)
Triangulum Australe	16S6A	16h 10m ± 8m	−60¾° ± 5′	(2)
Hercules	16N4A	16h 36m ± 10m	41° ± 2°	(13)
Hercules	16N0A	16h 45m ± 2m	6° ± 1°·5	(2)
Sagittarius	17S2A	17h 42m ± 1m	−28°·5 ± 0°·2	(21) May be associated with the galactic nucleus. The presence of neighbouring intense emission regions makes the measurements of flux density difficult
Ophiuchus	18S0A	18h 16m ± 4m	−8° ± 2°	(13)
Sagittarius	18S1A	18h 17m·9 ± 0m·2	−16° 14′ ± 5′	(18) Omega Nebula M17 NGC6618
Cygnus	19N4A	19h 57m 44s ± 2½s	40° 35′ ± 1½′	(3) Colliding galaxies
Cygnus	20N4A	20h 22m	40°	(11) Cyg X extended source possibly associated with galactic nebulosity
Cassiopeia	23N5A	23h 21m 36s ± 30s	58° 35′ ± 10′	(11) Galactic nebulosity

individual. Care, however, must be taken that exact multiples of the mains frequency are avoided. The values shown in Fig. 76 will give a frequency which lies in the region of 900 cycles to 1,000 cycles. This is connected to the drive unit because the output from the switch generator is not of itself sufficient to operate the phase switch and phase detector. The drive unit is therefore an amplifier of low impedance and the voltage is balanced about the earth point at the terminals 1, 2, 3.

An alternative circuit for the phase switch is shown in Fig. 77. This circuit was developed by J. J. Davies and has worked very well over a number of years. Here valves are used instead of diodes and the half-wave section is added between the two grids. The third valve helps to level up the amplification of the two sections. This unit is comparatively easy to adjust but careful matching of the stages is necessary.

The complete equipment has now been described which will enable the amateur to enter this exciting field of radio astronomy. A list of radio sources that can be observed is shown in Table 4. The Sun offers an immediate starting point for the amateur radio astronomer, and it is suggested that study of the Sun is the best way to become acquainted with the technique of radio astronomy.

Chapter Eleven

EXTRA-TERRESTRIAL RADIATIONS

WHEN WE LOOK up into the sky we see the bright section of the Milky Way. This concentration of stars indicates that we are looking towards the edge of our own star system, which we call the "galaxy". Containing as it does myriads of stars, it is an island universe in the immensity of space.

As mentioned earlier, our galaxy is of spiral form and has a large diameter and a somewhat thin section. Its extent is about one hundred thousand light years across and it has a thickness, at the point of greatest concentration, of some 20,000 light years. Our own solar system is at a point some 25,000 light years from the centre.

In 1944 Grote Reber had completed his radio map of the sky. His results also showed the maximum was in the direction of the Milky Way. This stimulated research for isolated radio sources in the background of the galactic noise. Hey, Parsons and Phillips made a survey of the sky at a frequency of 64 megacycles. Their observations also showed that radio emission was concentrated along the galactic equator. However, there was some difference between the results which were obtained by Hey and his team, as compared with Reber's results. Reber's results showed a background of steady activity, whereas Hey and his team found that these bursts were irregular. Hey therefore, quite rightly, assumed that this was something distinct from the general background which had previously been studied. Bolton and Stanley in Sydney, two other workers who made observations in this field, confirmed the suggestion. The conclusion was drawn that there existed a class of objects distinguishable from the background and which came to be known as "radio sources".

Many hundreds of these sources have since been mapped. Only a small number of them, however, have so far been identified with optical objects. Generally speaking, the size of the sources is rather greater than that of the size of the stars.

Indeed, there can be gas clouds which radiate and extend to several light years across. Certain difficulties face the radio astronomer in attempting to map the brightness of the radio sky. A highly directive system that this kind of observation requires, together with good resolution and high sensitivity is difficult to obtain and various methods have been adopted to reduce the size of the beam. Typical among these are the Mills Cross and other similar aerials. The system of aperture synthesis devised by Professor Ryle has proved to be most successful in this kind of work. Small aerial systems are not suitable for mapping sources of small size; however, with the instruments that have been used it is clear that the background radiations come from our own galaxy.

The Milky Way is always prominent along the centre of such radio maps, the brightness falls away as the distance from the galactic equator increases. An important point is, however, that it is not so intense in a direction at right-angles to the galactic equator. The contour lines show that the main structure of the background radiation is parallel to the equator and thus related to the structure of the galaxy itself.

From the point of view of the radio measurements our galaxy appears in the form of a sphere of radiation, the maximum coming from a direction towards the centre of the galaxy and a minimum in the opposite direction. It would appear, therefore, that these radio waves are generated in what must be a sort of halo which, however, is quite invisible optically.

The origin of these radiations is believed to derive from high-energy electrons. The mechanism of this operation is termed "synchrotron radiation". It was first described by Schwinger. Schwinger was not an astronomer, nor was he really concerned with radio waves. His particular interest lay in the activity of electrons inside accelerating machines where nuclear particles were taken to high energies. A synchrotron is an apparatus which whirls electrons around a circular track to which they are confined by a strong magnetic field. At high energies the electron beam which is produced emits a blue glow of light called "synchrotron radiation". This radiation was shown by Schwinger to follow naturally from the acceleration of electrons in a strong magnetic field. Analysis of this radiation shows that

it varies with electron energy and also with the strength of the magnetic field which bends it in its path.

The magnetic fields which we encounter in interstellar dust are only a small fraction of the very high order of magnetic fields which are used in synchrotrons. However, the areas in which electrons exist in our galaxy are so great that even though the electrons are far apart there are many of them. A small concentration of electrons at high energy can therefore account for the total radiation that we find in the halo.

There are a number of definite sources which have been identified with objects in our own galaxy. We find that some of these are the remnants of supernova explosions. One particular source is that of the Crab Nebula. This exploding star was witnessed by the Japanese and Chinese astronomers about the year 1054. A star which had previously been invisible suddenly exploded and blazed up as bright as Venus. Naturally, it attracted immediate attention. After a time it faded away again until it could no longer be seen with the naked eye. Since the time of the event, this great turbulent mass of gas has been expanding and rushing out into space at the rate of 700 miles a second. Its present size now approaches a diameter of some ten light years. This means that it is so large that if one edge were in contact with the Earth the other would reach out as far as the bright star Sirius in the constellation of Canis Major.

The Crab Nebula was discovered again as a faint object in the nineteenth century. The name given to it is the result of the contorted appearance of the mass of gas with a bright centre section. The radio observations have shown that the spectrum is rather different from that of other sources. The amount of energy reaching the Earth is almost the same at all frequencies and for this reason it was at first thought that the Crab Nebula must be giving rise to radiations from ionized gases of the type known as "black-body" emission. However, it was realized that the mass of the nebula was not sufficient to support so great an output of energy as would be required by the black-body theory.

The Russian astrophysicist, Shklovsky, suggested that both the radio and optical emission might be due to synchrotron effects. It is possible to identify synchrotron emission because it is polarized. At the observatory in the Crimea Drombrowski set up an experiment to determine whether, in fact, the radiation was polarized. In this he was successful, and his observations showed that the synchrotron-type of emission was the means of the production of energy from this source.

J. H. Oort and Walraven of Leiden University concentrated on optical studies of the Crab Nebula. In America Baade at Mount Palomar was able to show that the magnetic field was closely associated with the filaments of the network which appear to surround the nucleus. These filaments took up a position very similar to that which iron filings do when they are dusted on to a piece of cardboard and held over a bar magnet. It was observed that ripples of light moved outward from the central nucleus at a speed of one-tenth of the velocity of light. Further attempts were made to detect polarized radio emission from the Nebula. At Jodrell Bank and again at Leiden these measurements were made and the polarization was found to be less than a few per cent, which is much lower than the optical figure. Since, however, the radio source is twice the size of the optical Nebula, and the radio waves appear to come mostly from the outer parts, this was not surprising.

Another point that has to be remembered is that Faraday rotation is present and this would account for the anomalies in polarization. Later measurements confirmed, however, that polarization was present and these were made at very high frequencies where the Faraday effect is not so great.

The energy necessary to create the magnetic field in the Crab Nebula is thousands of times the amount that our Sun would have to put out during the 900 years since this cosmic explosion was seen. The only explanation is that such a source of energy can only come from radio-active decay. From the observations, both optical and radio, it was clear that if the emission were caused by cosmic-ray electrons most of the energy would have been radiated away in less than 300 years. The Nebula is, however, 900 years old and it is still possible to observe it by its own emitted light, so this must mean that new cosmic rays are being created continuously. It may well be that the same source of energy that created the magnetic field is also responsible for these other effects.

Other supernova explosions have been identified and one such was observed in 1572; this was called Tycho Brahe's star. There is a wisp of gas observed within half a degree of the reported position of this star, and this probably is a remnant of this explosion. Another was observed by Kepler in 1604. The remnant of this has been observed, although it is very much fainter than the Crab Nebula.

A number of other sources which have been identified are very large, almost circular clouds of gas. They show filaments of highly activated gas which are expanding rapidly. One of these is the Cygnus Nebula, sometimes called the Veil Nebula. It is an almost circular streamer of gas about 3 degrees across. This also has a similar level of radiation extending over a wide range of frequencies from 100 megacycles to 1,200 megacycles. A spectrum of this nature is almost certain to indicate that the radiations are of thermal origin produced in a very rarefied and ionized gas.

Another thermal radiating cloud of ionized gas is the Lagoon Nebula. This is some 3,600 light years distant from us and is about fifty light years across.

There is a remarkable radio source known as Cassiopeia A. At the Cambridge Radio Astronomy Observatory a special radio telescope was built in 1948 to study the newly-found Cygnus A source which had previously been reported by Hey in England and Bolton in Australia. After observations had been made the records were examined and in addition to the Cygnus A source there was another source nearby which completely outshone Cygnus A. Bolton had not seen it because it was not visible in Australia, being too far to the north. Hey and his group had not detected it for the aerial system they used looked only out towards the horizon. No obvious identification was made by the Cambridge workers at the time, but F. Graham Smith was given the task of plotting the accurate position of this new source. He decided to identify it, if possible, with a visible nebula or star. Another telescope was built and put into operation at a frequency of 214 megacycles and with this equipment the position of the new source was accurately found to within ten seconds of arc. It was given the name Cassiopeia A.

The facts were then passed on to the optical astronomers in Cambridge, and Dewhirst was able to show that there was a very faint nebula near to the position that Smith had allocated to Cassiopeia A. At Mount Palomar Baade and Minkowski were asked to take steps to check the position with the 200-in. optical telescope. It was felt that such a powerful source should be easily visible, but this was not found to be the case. The object identified by Dewhirst proved to be a small block of nebular gas. The actual radio source was distributed over a wide area of five minutes of arc in diameter. All over this area were faint clouds of gas. This is all that remains of the original hot sphere which must have been the result of the original supernova.

On examination it proved to be a most interesting object. Parts of it were found to be stationary and parts were moving at terrific speeds of 3,000 miles a second. At this very high speed it was possible to see changes even over a short period of a few months. Photographs taken at long intervals of a year or so have shown these changes to be so great as to render recognition almost impossible.

This is another instance of atoms having been stripped of their electrons by thermal agitation. The highly ionized gas is the sort of thing that we know to generate radio waves. This again is no doubt an example of synchrotron radiation. From the speed of the movements of these small bits of clouds, calculations have been made which show that the actual explosion must have dated back to around the period 1702. As it is some 10,000 light years away from us it is not surprising that this explosion was not noticed, for at this distance, even at the peak of its explosion, it would not have been bright enough to be seen by the naked eye.

Another radio source is found in the constellation of Auriga. Two smaller sources which are nebulosities appear in the constellations of Puppis and Gemini. The measurement of these is very difficult because they are of a rather large angular size. The process by which they emit radio waves is not fully understood and they do not seem to be increasing in size very rapidly, but it is possible that they may be remnants of exploding stars that are very much older than those that have already been dealt with. The gas may now be so diffuse that the original expansion is now halted by mingling with other clouds of interstellar gas.

Another source of radio waves is those generated in clouds of ionized hydrogen. This is the kind of cloud which surrounds some of the hottest stars of which an example is the nebulosity around the central hot star in the Sword of Orion. This object is visible to the naked eye on a clear night. A pair of binoculars or a low powered telescope enables us to discover that it is in fact a nebula. It appears as a small cloud rather than as a distinct point like a star. This is probably the kind of thing that stars are made of. The nebula contains many young stars which are extremely hot and it would appear that they are streaming away in opposite directions at very high velocities, sometimes of the order of 150 miles a second. In this region there are nearly a hundred stars.

In contrast to hot stars where hydrogen is ionized at a temperature of some 10,000° Kelvin, like those of the Orion Nebula, we have cold hydrogen. Far away from the stars where there is nothing to keep the gas at a high temperature the level falls to about 100° Kelvin. This is where we find the hydrogen line. This spectral line of radio emission has been dealt with under the history of radio astronomy. The temperature can be deduced from the intensity of the line just in the same way as the temperature of the hot stars in the Orion Nebula was deduced from the intensity of the radiation of the ionized hydrogen. The study of this spectral line shows that our galaxy has the same spiral formation of many of the common galaxies in the universe. The profiles of the spectral line show that our solar system is in one of the spiral arms. This arm is called the "Orion arm".

There is another expanding arm near the centre of the galaxy, another in Sagittarius and still another in the region of Perseus.

There are some hundred thousand million stars in our galaxy. From the optical point of view the galaxy consists of two principle parts, for in the main it contains type-2 population stars, and these probably account for the bulk of the galaxy. The more spectacular parts, such as the spiral structure, contain type-1 population stars. It is often the case that radio astronomy brings us information about what lies between the stars rather than that of the stars themselves. Radio emission comes from the very cold hydrogen at 21 centimetres and from the hot clouds that surround the newly-formed stars.

Observation of this emission enables us to decide that part of our galaxy is moving in one direction and part in another.

There is some difficulty in attempting to form a picture of our galaxy while we ourselves are within it. It forces upon us the necessity of making a two-dimensional picture of something that completely surrounds us. However, we are able to study a very similar galaxy to ours and one which, fortunately for astronomers, is close enough to be examined in detail. This is the great nebula M31 in Andromeda.

Extra Galactic Nebulæ

Immediately after World War I our knowledge of the universe outside our own galaxy was extended enormously. This came about when the great 100-in. telescope at Mount Wilson in America was completed. This instrument, which had more than double the light-gathering power of any previous telescope, was put to work by the astronomer Edwin Hubble. He devoted the majority of his attention to the problems of the distant nebulæ. Of all the astronomers of recent years Hubble must rank as one of the greatest, for he revolutionized our ideas of the universe by showing that stars were in collections or groups. He showed that our own galaxy was only one of the many island universes, some of which were so far away as to appear but small faint clouds. Many of the galaxies that we know are of the spiral type. Existing at enormous distances from our own galaxy the nearest of these in our family of galaxies is the Andromeda Nebula.

The Andromeda Nebula is the only galaxy that can be seen with the naked eye from Europe and on a very clear night a faint hazy spot may be seen in the constellation of Andromeda. With a pair of binoculars, or a low-powered telescope, the general shape can be observed. It is our nearest large neighbour but it is at a distance of some 2,000,000 light years. The diameter of our own galaxy is 100,000 light years. Our nearest neighbour, therefore, is more than ten times as far away as it is from one side of our galaxy to the other. As astronomical distances go this makes the Andromeda Nebula our next-door neighbour. Astronomers are, therefore, fortunate in being able to study in detail the structure which is very similar to our own. From our

own position we see the galaxy at a slight angle, and it is for this reason that, although in fact it is circular, it appears elliptical.

Powerful telescopes reveal bright stars in the spiral arms. Some of these, like the stars in the Orion Nebula, have shells of ionized gas around them. Were we able to place ourselves inside this galaxy we would find that it too, had a Milky Way, somewhat similar to our own.

The Andromeda Nebula also has a halo of radio emission. A radio astronomer is therefore able to study at close range a common type of galaxy and deduce information about our own galaxy from the study of our neighbour. Sir Bernard Lovell and his team at Jodrell Bank were the first to detect radio waves from the Andromeda Nebula. This was before the 250 ft radio telescope was in operation. After this had come into use a more detailed exploration of the Nebula was made.

The Cambridge workers have also examined the Andromeda Nebula in detail. From the observations made it is now quite certain that the radio galaxy is very much larger than the visible galaxy. It appears to be spherical in shape and this agrees very well with the discoveries about our own galaxy. Among other things, it has shown that the Andromeda Nebula has a magnetic field. The extent of the visual size is about 100,000 light years, but the radio halo is about twice this size.

At Leiden University astronomers have studied the spiral arms. Using their 21-centimetre line receiver they were able to study very carefully the shape and size of the arms. Observations have shown that the maximum speed of rotation is of the order of 150 miles a second.

The output from this galaxy is about the same as that from our own. This is one of the reasons why the first attempts to detect radiation failed, for they were obscured by the radiations from our own galaxy.

In the southern hemisphere there are two additional galaxies visible which are also near neighbours. These are the greater and the lesser Magellenic Clouds. They are irregular in shape and one is very much larger than the other. Optical telescopes enable us to resolve the individual stars quite easily and these provide us with a useful measurement of astronomical distances. The two clouds are joined together by a tenuous band of gas and gravitational attraction between them causes the clouds to rotate about one another. Compared with other galaxies they are untidy clusters of stars.

From the observations made it would appear that there is more dust than gas in this irregular type of nebula than those which have developed into spirals. However, while the large cloud contains much cosmic dust, very little has been detected in the smaller cloud. Using the 21-centimetre line, radio astronomers have investigated the clouds. Part appear to move at different velocities of rotation and the maximum is shown to be in the region of 12 miles a second, which is very much slower than that of our own galaxy. One suggestion is that the hydrogen gas clouds are much younger than those in our own galaxy and that only some of the cloud is condensed into stars. Both the clouds are moving away from us and, in the case of the larger cloud, the speed is of the order of 200 miles a second, and in the case of the smaller cloud at about 100 miles a second. de Vaucouleurs has claimed that the upper portion of the large cloud exhibits a spiral structure, and he has also suggested that a tenuous bridge of gas extends from the large cloud to our own galaxy. Close examination of the 21-centimetre line maps which have been made of these galaxies shows there would appear to be a corona of rarefied hydrogen surrounding them both.

Turning our attention to much greater distances from our own galaxy we find many clusters of galaxies. The 200-in. Mount Palomar optical telescope is able to measure nebulæ of the Andromeda type up to a distance of 1,000 million light years. There are probably more than 100 million of such groups of galaxies and their distribution in space is a matter of considerable importance to cosmologists.

Hubble has taken many photographs of these extra-galactic nebulæ. He found that they vary in shape and size, some have wide open spiral arms, some have tightly closed arms, and some are twisted into curious shapes. One type is known as the barred spiral. Others are almost spherical and some are very flat. They all exhibit characteristics which show them to be a mixture of the various types. Of these, some have been designated "peculiar" and others have been found to be galaxies in collision, or a state of separation.

One such object in the constellation of Centaurus proved to be 2 degrees across. This is distant from us about 100 million light years. The radiation is so great that it was supposed that the whole group was radiating. This set a problem for radio astronomers because the object radiated about four times as much as would normally be expected from a group of its size. It was, however, pointed out that there was a "peculiar" galaxy in the centre of the object that was possibly the radio source. Using an interferometer, detailed observations were made and it was discovered that this was, in fact, the case. Seventy-five per cent of all the radiation came from one nebula which was identified as two galaxies in collision. An enormously high level of radiation, which surprised the radio astronomers, was therefore explained. This is one of the strongest sources that has been discovered from the southern hemisphere. Optically it appears almost spherical and has a dark band stretching across it and it would seem that this dark band is a flatter spiral galaxy which is mixed up with the spherical nebula. The dark band appears to rotate, and this is added evidence that the whole system is in fact a collision.

The first accurate position of an extra-galactic source was Cygnus A, and this work was carried out by F. G. Smith at Cambridge. The actual identification of the source proved to be extremely difficult. Eventually from photographs that were obtained with the 200-in. telescope at Mount Palomar, identification was possible. The source was found to be two galaxies in violent collision. Parts of it were flying about in all directions at high speeds up to 300 miles a second. The size of the source observed with the radio telescope appeared to be many times the size of the optical object. It also seemed that it was separated into two distinct parts, one on either side of the visible object. It would appear that when two galaxies collide the dust and gas contained in them is slowed up by the collision. The original energy stored in the motion of each of them will subsequently be converted into thermal energy. This will cause the particles to heat up to a very high temperature and to radiate. The amount of energy radiated from this collision is calculated to be equal to one hundred thousand million suns. It is a fortunate fact for us on Earth that it is so far away, for were it, for example, as close as the Andromeda Nebula, radio communications would be impossible. Even with the protection of the atmosphere it is hardly likely that any living thing could exist on the Earth for the atmosphere would be bombarded with so much energy that it would glow with a perpetual aurora.

There are a number of definite cases of collision among the identified sources. There is, however, only one which approaches the condition that we have found in Cygnus A. This is the radio source of Perseus. This galaxy is much closer to us than Cygnus A, being at a distance of some 170 million light years. This is a most energetic collision between two nebulæ. One of them is a tightly wound spiral, while the other is a much more open type. The one which is open shows considerable distortion from its normal shape, and from the spectrum we have been able to study it is found that they are so distant that they are both moving away from us at terrific velocities. One appears to be moving at about 3,000 miles a second while the other is moving at about 5,000 miles a second. The difference in this speed is, in fact, the velocity of the collision itself and is about one per cent of the velocity of light. At such a tremendous speed all signs of collision would be over in about a million years. The size of this great cosmological catastrophe is about the same as that of our own galaxy, 100,000 light years.

There is another "peculiar" galaxy in Virgo A, which proves to be almost a sphere. The diameter of the radio source appears to be the same in this case as that of the optical object. This is a "peculiar" galaxy and not a case of collision, and its peculiarity is that it has a bright blue streak projecting out from the centre. Just as the light from the Crab Nebula was polarized, so is the light from this jet.

These are the only two cases known and it is certain that the origin of the radiation is from the same cause, that of synchrotron emission due to cosmic ray electrons in the magnetic field.

In the case of Virgo A, however, a much greater explosion has occurred than that which occurred in the Crab Nebula. The mass of Virgo A is about a million times that of the Sun.

The most distant galaxy that has so far been identified is in the constellation of Boötes, which is situated at a distance from us of some 5,000 million light years. This is a galaxy that would probably never have been found without the aid of radio astronomy.

About a hundred of the many thousands of radio galaxies have actually been identified with optical objects. One of the important reasons for studying these far-distant radio sources is to determine their distribution throughout space. From the point of view of cosmology it is important to know whether they are concentrated in groups near our galaxy or more concentrated at very great distances. Cosmology, which is the study of the whole universe, seeks to formulate some definite explanation of the origin and the size of the universe. Radio astronomy has become an important new tool in this respect, for the limit of our knowledge of the universe is restricted by the instruments used for its study. Radio astronomy can extend observations out into the depths of space far beyond that which the optical telescope can penetrate.

At the moment it is thought that the limit will be some 10,000 million light years. The reason for this is that since electromagnetic waves move at the speed of light, at this vast distance the nebulæ, which appear to be receding from us, would in fact disappear from view as they would be moving away from us at the speed of light, no radiation could reach us. If, however, the distance is less than 10,000 million light years, we should be able to receive some of the radiations. We can observe whether nebulæ are receding from us by the amount by which the radiations move from a higher to a lower level of frequency in the spectrum. This is the "red shift". At the moment the only satisfactory way of doing this is by observation of the normal visible spectrum.

It is necessary to remember that when looking at anything at these vast distances they are, in fact, being viewed in the past. In the case of the nebula which was found to be 5,000 million light years away, it is seen as it was even before the Earth was formed. Whatever model we make of the universe in which we find ourselves, we can only picture it as it was at this vast distance back in time.

There are two main theories of the universe. The evolutionary theory and the steady-state theory. For the evolutionary theory it is stated that the universe started about 100,000 million years ago with a tremendous explosion. It implies that all the matter of the universe was in a very condensed state, one small lump as it were. Observations that have been made of galaxies receding at terrific speeds are said to be the result of a dispersal of matter from this original explosion. The galaxies moving fastest will have gone furthest out into space. If an astronomer studies these galaxies which are at vast distances it should be possible for him to look at them as they were much nearer to the time when the explosion occurred. Those who believe in the steady-state theory say that the universe is the same all the time, if it is studied carefully. It is said that the galaxies are formed, grow up and gradually die and disappear completely. This is supposed to be going on continually. The galaxies receding at vast speeds reduce the density of the matter, this loss of matter is then counter-balanced by a continuous creation of hydrogen atoms. The rate of creation that is required to do this is extremely small, so small in fact, as to be undetectable by any method which is available to us at present.

Recently Professor Ryle, when giving the results of some years of measurement of these far-distant galaxies, suggested that they were concentrated at these vast distances. His results, he claimed, are highly suggestive of the evolutionary, or exploding universe. Though it is too early to form any positive conclusions, it nevertheless now requires that a new look should be taken at the steady-state theory. The evolutionary theory predicts that galaxies in collision were, at a time long ago, much commoner than they are now. The steady-state theory on the other hand, predicts that the large-scale of the universe is such that over very long time-periods involved it does not change at all.

Radio astronomy may well provide the answer to these problems in the end.

The Moon

Radio astronomers have studied the Moon very carefully over many years. A great deal is known about its surface structure and there are extremely detailed maps of its surface. However, even with the best methods of optical observation, there is still very much that we would like to know about it. Because the Moon makes only one revolution on its axis as it goes round the Earth, we are normally only able to see the side which is presented to us, and this is always the same. The lunar probe, Lunik III, put up by Russia, provided us with pictures of the other side of the Moon. Although these pictures were crude, they served to show that the Moon is generally very similar all over.

At the frequencies used in radio astronomy the reflected radiation due to sunlight is very small. We are thus able to study radiations emitted by the body itself. Predictions can be made as to the level of this radiation. However, when the actual measurements are made it is shown that they are somewhat less than the predicted figures, one reason being that the Moon is only a partial radiator. Observations which are carried out at the time of eclipses show that, although reflected sunlight was cut off during the period of eclipse, the radio emission remained more or less constant. It would have been expected that if the radiations were coming from the surface of the Moon there would have been a considerable reduction when the sunlight was cut off. It does not act in this way and therefore this is an indication that beneath the surface of the Moon the temperature remains more or less constant. The depth is shown to be a little over 2 in. at the shortest wavelength, at which the measurements were made. When the heat of the Sun was cut off the temperature of the surface fell very rapidly. We can suppose from this fact that changes of temperature do not penetrate to this depth of two inches and that the Moon's surface is a good insulator. If this surface was solid rock this kind of insulation could not exist. If, however, it consists of dust or small particles, then we might have an explanation of the way in which the radio waves behave. It is not possible to tell precisely how deep this dust is, but it cannot be less than about 1 in. nor more than about 3 ft.

In 1946 radio waves were directed at the Moon and the echo of the returning signal was noted 2½ seconds later. A tremendous amount of power is needed to get such an echo back from the Moon. The same experiment was carried out in Australia by Tower and Shain, who did it with one of the Australian radio stations. The beam of this station at a frequency of 31 megacycles was directed at the Moon, and the echoes received back changed very rapidly in strength and at times could not be heard at all. There was also slow fading of the signal and the astronomers decided that this fading was due to the wobbling of the Moon. This "libration", as the act of wobbling is called, is due to the fact that the Moon presents different aspects to the Earth. When the Moon rises, the observer is on one side of the Earth; later, when the Earth has moved round on its axis, the observer has moved to the opposite side of the Earth and the Moon is setting. Since the Earth's axis is not perpendicular to the plane of the Moon's orbit, and also because the Moon's orbit is not quite circular, the radio waves will be reflected from different parts of the Moon's surface. The irregular surface of the Moon will tend to produce echoes varying in strength and sometimes fade very rapidly.

The same experiment was carried out at Jodrell Bank by J. B. Evans, who confirmed the work of the Australian radio astronomers. He was able to show that at one particular frequency the waves were reflected only from a central part of the Moon. It was apparent that this particular area appeared of different size to different frequencies of radio waves. At 200 megacycles this area proved to be 200 miles across, but when the frequency was raised to 3,000 megacycles it seemed to be almost a perfect mirror, but only 35 miles across.

Since it is such a good reflector, the Moon could be used as a relay station, that is to say it would be possible to project signals to the Moon from one side of the Earth, which would then reflect them to the other side of the Earth. Since 1954 this has been done many times. It is possible to measure the distance of the Moon from the Earth by means of radio waves. A radar pulse sent out to the Moon and back again will take about 2½ seconds. By this method the distance between the Earth and the Moon has been computed to an accuracy of half a mile. The Earth is irregular in shape and creates certain errors in these measurements. These very errors are a means of determining the shape, or figure, of the Earth.

The Planets

Thermal radiation has been detected from a number of planets. During the closest approach by Venus to the Earth in 1956 American radio astronomers paid special attention to the planet. Using wavelengths between 3 and 9 centimetres and a 50-ft diameter dish, the brightness temperature was found to be about that of the melting point of lead, that is 300° C. This was an unexpectedly high level, since it was the dark side of the planet which was being observed. Using still shorter wavelengths observations have shown that this level of temperature falls off considerably at the higher frequencies. This would indicate that the high frequency waves come from the cooler regions surrounding the planet. Though Venus is much closer to the Sun than the Earth, it is still not close enough to enable the Sun to raise the temperature to such a high degree as that observed by radio; there may be some other explanation for this. Since Venus appears to be covered in cloud it may very well be that these clouds prevent the heat from getting out.

Various attempts have been made to get radio echoes from Venus. The distance of Venus from the Earth measured by this method varies considerably, but the last attempt, which was made by Jodrell Bank, would appear to be most accurate. This experiment of sounding the surface of Venus is very important to astronomers. The actual surface of the planet is never visible, for it appears to be covered in dense cloud. We therefore know nothing whatever of its surface. Radio methods may well enable us to determine the length of the period Venus takes to rotate on its axis; until now more than eighty different measurements have been made, which range from a few hours to many days.

Radiations from Mars were detected in 1956. The brightness temperature proved to be quite low, but near to the calculated value that would be expected from thermal radiation from the surface of the planet. The power required to get an echo from Mars would be many times greater than that required to get an echo from Venus. Even when Mars is closest to the Earth, which will next be in 1971, the power required to get an echo from the planet will be ten times greater than that required to get an echo from Venus. There are other difficulties involved also, for it would take 7½ minutes for the journey of the signal when dealing with Mars.

Thermal radiation from Jupiter was also in close agreement with other methods of detection. Jupiter, however, is also responsible for radiations which are of non-thermal origin. In 1955 Burke and Franklin observed outbursts of noise which were extremely intense. This takes place within a limited range of frequencies, between 18 and 26 megacycles. The activity goes on for as long as one or two hours and the radiations take the form of groups of intense bursts, some of which last about a second, but many are of much shorter duration. The strength of these outbursts considerably exceed those of the most intense sources. The reason for this radiation is not understood. Sufficient is known about Jupiter itself to be able to say that a thermal origin is quite impossible. One of the early suggestions was that it was due to thunderstorm activity, but the activity of thunderstorms and lightning storms is understood on Earth. If such a mechanism existed on Jupiter it would have to be very different from that which we observe on Earth.

Another peculiar feature of these radiations is that they appear to show circular polarization. Though they are confined to the frequency band between 18 megacycles and 20 megacycles, outbursts do not occur simultaneously all over the band. It is sometimes observed at two or more frequencies at the same time, or it may be observed at one frequency and not at another. The most frequent bursts appear to be in the region between 18 megacycles and 20 megacycles. After the original discovery by Burke and Franklin, Shain in Australia examined some of his early records at a frequency of 18 megacycles. These records had been made during the testing of the Mills Cross in Australia and were not devoted specially to the study of Jupiter. The records, however, showed noise peaks comparable with those discovered by Burke and Franklin. These observations in Australia had been made during 1951. In 1954 Burke and F. G. Smith showed that there was a predominence of outbursts at 22 megacycles; in 1955, 22 megacycles was again observed. Barrow, Carr and Smith carried out extensive observations in 1957 and at the present time Barrow is concentrating on work in this field.

Attempts have been made to correlate outbursts of noise with optical features on Jupiter. No success so far has been apparent, although it was thought that there might be some correspondence between the radiations and the Red Spot on Jupiter.

Polarization measurements are such as to indicate that no contribution to this is made by our own ionosphere. Barrow has suggested that there is an ionospheric layer on Jupiter of the same nature as the Earth, and from this he has made a tentative suggestion as to the electron density. The fact that the waves are polarized implies that there must be a magnetic field associated with Jupiter, and it is probable that some, at least, of this radiation is due to plasma oscillation in Jupiter's upper atmosphere. There is sufficient ionization in Jupiter's atmosphere for this to take place as a result of solar radiation.

Thermal radiation has been detected from Saturn and this radiation again agrees with that expected. It has been suggested by workers in America that non-thermal radiation has also been detected from Saturn, but this has not been confirmed.

Chapter Twelve

INTO PRACTICE

HAVING NOW COLLECTED together the information which we will require for the construction of a radio telescope it is necessary to consider some plan for its use.

In order to devise a programme it is first necessary to decide the type and size of radio telescope that will be built. There are a number of factors which determine the choice; the first of these is the space that is available for the erection of the aerial. The average size of a garden will be somewhere between 60 and 100 ft in length and 40 to 100 ft in width. This provides enough area in which to work, but the proximity of other buildings, power lines and trees must be taken into account. This point is particularly important where the high frequencies are concerned. A simple arrangement of either the Kooman array or the corner reflector on a rotating stand requires a reasonably clear view unobstructed by buildings for at least 100 yd in the southerly direction. If this is not possible and there is a clearway at some other point of the compass, then only those radio sources which are circumpolar will come within the programme of study. If the area, therefore, is severely restricted and the sky is open only to the zenith, as would be the case if the garden was surrounded by trees or by tall buildings, or other obstruction, it will be possible to study the intense sources of Cygnus and Cassiopeia by means of an aerial which may be permanently directed vertically upward. In this case, if the aerial is intended to be a more or less permanent fixture, three choices are available. First, there is the fixed Kooman array which may be arranged on one of the single standard sections; or an alternative is to erect the reflector above normal head level, say 7 or 8 ft, and extend this across the whole width of the garden. Such an arrangement if 30 ft long by some 10 ft wide will provide an aerial of a comparatively narrow polar diagram in one direction. This would be ideally suited for the examination of sources within several degrees

north or south of the vertical position. Such an arrangement would provide a natural archway and not interfere with the activities carried on in the garden. It also has the advantage that a wide range of frequencies could be used for observation, since the length would permit the longer elements of aerials required for the lower frequencies. If it is possible to arrange such an aerial east and west the beamwidth at 200 megacycles would be of the order of 10 to 12 degrees. The gain of such an aerial would be very high and isophotes of a strip of the sky containing Cygnus and Cassiopeia would be possible.

If the corner reflector is used the same principle applies, except that it would be more useful if made steerable in altitude. With the comparatively narrow beamwidth of the V it would be possible to scan successive strips within the limits of the visible portion of the sky. Even under the severest conditions of trees or pylons or other buildings surrounding the garden some 30 to 40 degrees of sky could be covered.

Another alternative in the fixed type of aerial is the multi-element Yagi. The gain of this, of course, is higher with a smaller aerial, but the frequency is restricted to within a narrow band on either side of the centre frequency for which the Yagis have been designed.

It will be clear then that even under the most severe restrictions something can still be done. Where tall blocks of flats are involved and the amateur has access to the roof which is flat, it is possible for interferometers to be mounted with a considerable increase in the versatility of the station. This also applies to schools where it might be preferable to have the aerials on the roof rather than occupy sections of the playing field or playground. The aerials themselves can be made portable so that they can be set up for an experiment and then dismantled and put away. This, though a tedious business, might be preferable to having nothing at all if there is an objection to having an aerial in a fixed position.

In the case of simple radiometers the receiver will be quite straightforward. If the receiver proper is located at some distance from the aerial itself it might be preferable to use a pre-amplifier at the aerial. Transistor-operated amplifiers are available commercially at quite reasonable prices and this, in association with low loss cable, will provide a reasonable signal at the receiver. If the gain of the amplifier is of the order of 20 db it will compensate for the loss of the cable up to 200 or 300 yd.

The choice of the receiver depends upon the funds available. Already in Chapter 8 a number of receivers have been suggested, both for direct use on the aerial and as intermediate amplifiers. Apparatus of this kind is of very high-class workmanship and extremely reliable. Where funds are available there are many commercial types of communication receiver which would form an ideal basic unit for this purpose. However, one of the principal reasons for amateur activity is because an amateur prefers to make his own apparatus and the details, which have already been given in Chapter 10, enable a start to be made.

Readers who are themselves electronic engineers or are able to call upon a friend who is an electronic engineer will find many types of receiving circuits available in the literature which could be modified or adapted for the purpose of radio astronomy.

Amateurs fortunate enough to have a space to erect an interferometer will be able to extend their activities very considerably. It would be best to start with the simple drift interferometer and having gained some experience with this the more sophisticated phase-switching interferometer could be put into operation.

If a pen recorder is available then the apparatus can be allowed to work automatically even when the observer is not present. The recordings which have been made can be examined at leisure and then necessary interpretations made. When no pen recorder is available the two alternative systems of tape recording or translator, or the two used together, may be adopted. If no tape recorder is available then the observer will have to be present when events are taking place.

Whichever system is used it will be necessary to set up a point of calibration. This will be done with the noise generator applied to the input of the receiver. In the case of the pen recorder this need only be done at the beginning and the end of the recording period. In the case of the translator the noise should be set to a predetermined level and this level noted in the log book. Readings should then be made at subsequent

intervals depending upon the size of the source which is being received and the beamwidth of the aerial. If the beamwidth is comparatively narrow, say 10 to 15 degrees, then measurements should be noted once every two seconds. If the source is more extended or the beam of the aerial wider, then longer intervals will be satisfactory. It is convenient to listen to the sound in the monitor speaker at the same time as taking the meter readings. This will enable the observer to become familiar with the different types of interference that appear; due allowance can then be made reading the pen recordings or properly interpreting the tape recordings.

After a few weeks of operation the observer will become quite familiar with the many and varied types of noise which appear. This, like any other subject in astronomy, is a case of practice leading to greater perfection.

What has now been described is the simplest of all the arrangements. Moving on to the more sophisticated phase-switching interferometer, enough has already been said to indicate the necessity of having very stable supplies and apparatus which is in first-class condition. We must here assume that the facilities are available for erecting an interferometer. The actual size and distance apart of the aerials will be determined by the space available, but the decision as to which frequency at which to work will depend upon the observations to be made. If small sources are to be observed then the higher frequencies will be required in order to get the longest possible base-line. Where the base-line available is more than 50 wavelengths one of the observations that can be undertaken will be the occultation of radio sources by the Moon. This will, of course, require extremely sensitive receivers, for it is quite difficult to discriminate between noise and actual occulted sources. At least 100 wavelength spacing is recommended if this kind of work is to be seriously attempted.

If a general observation of the sky is to be undertaken where sources may be large and irregular, the lower frequencies can be chosen with correspondingly shorter base-line.

Generally speaking it is advisable to make the maximum use of the space available even if this should be a few degrees north or south of the east/west line. Departure up to 20 degrees is still useful and the necessary qualification can be made of the records in order to allow for the passage of the source through the beam of the aerial.

One important point to remember is that the aerial system must be versatile, otherwise the variety of observations is likely to be restricted. As a general rule it is recommended that in the case of an interferometer at least one of the aerials should be steerable in azimuth as well as in altitude. It is possible to use one large fixed aerial and one small movable aerial. If there is space available, for example, to erect a corner reflector some four or more wavelengths long, the second half of the interferometer could well be one which is only two wavelengths long but has the full steering facilities. The overall result with the two aerials of different size will be less than two large aerials but not so much less that the records would be invalidated.

All the aerials used in any case should have some degree of steering in altitude. A very useful interferometer can be made up by utilizing pairs of Yagi aerials on a long base-line, and one such aerial is described by Osborne in a paper in the Memoir of the British Astronomical Association noted in the bibliography.

The practice of radio astronomy is essentially one which lends itself to teamwork as well as to individual observations. If this is combined also with a general interest in astronomy much useful information can be compiled; where two or more groups are situated at a distance from one another a number of observations may be undertaken simultaneously to determine the effects of scintillation. Thus, not only would the individual group have its own observations to make but co-operation with other groups could lead to greater knowledge and fruitful discussion.

A possible programme of work for a year is as follows. For the simple radiometer the most obvious object for study is the Sun. This observing programme should be carried on throughout the entire year, observations being made one hour before noon until one hour after noon G.M.T. Where restricted views of the horizon arise then the effect of the Sun on rising could be observed, or if the view were restricted to the west then observations at the time of its setting could be undertaken. If a steerable aerial is used then observations could be divided into three sections; those at sunrise, those at midday and those at sunset. Observations of the Crab Nebula, and in June of each year the occultation of the Crab Nebula by the Sun's corona, could also be undertaken. In this case it will be necessary to plot accurately the position of the Crab Nebula in relation to the radiometer and observe this and the Sun over a period of one month before and one month after the occultation. The gradual merging of the two sets of radiations can be noted, particularly at the time when the Crab Nebula ceases to be distinguishable from the effect of the Sun and again when the two separate.

Observations of Cassiopeia A are particularly valuable, for it is suspected that this source is decreasing in intensity annually. Observations of the other powerful sources indicated in Table 4 can be followed up in the same manner.

This programme can also be undertaken using the simple interferometer or the phase-switching interferometer. Considerable increase in position accuracy will be achieved by this method and also it will be possible to determine the actual strengths of the radiations received. Where the interferometer is large enough then occultation of radio sources by the Moon can be attempted.

Amateurs living in the southern hemisphere will find that some of the more intense sources are below their own horizon. There are, however, one or two sources which can be studied with the telescope already described (Table 4).

One very useful programme that can be undertaken is the study of that area which is contained in the Cambridge 3C Catalogue of radio sources and also the Australian measurements of the same area. These two surveys have never been completely reconciled. One advantage that amateurs in the southern hemisphere have over those in the British Isles is that generally they have more space available and can therefore erect interferometers with long base-lines. This will enable the necessary additional sensitivity to cope with the greater number of faint sources that are within their reach. Conditions of interferences are also different in the southern hemisphere and once again scintillations in the over-lapping areas of the northern and southern hemispheres can be profitably studied.

The southern hemisphere also offers better opportunities for work which can be carried out on the low frequency end of the radio spectrum; this could mean that useful work could be undertaken on the study of radiations from Jupiter.

It is important that observations are made as regularly as possible and a constant check made with members of other groups in order that the results may be compared and analysed. In this connection results can always be forwarded to the British Astronomical Association, Burlington House, Piccadilly, London, W.1. A special Section exists which deals particularly with radio astronomy; this is under the directorship of Mr J. Heywood, F.R.A.S.

BIBLIOGRAPHY

GENERAL ASTRONOMY

LYTTLETON, R.A. *The Modern Universe*. Hodder & Stoughton, 1956.

MOORE, PATRICK. *Astronomy*. Oldbourne, 1961.

MOORE, PATRICK. *The Amateur Astronomer*. Lutterworth Press (4th Revised edition, 1961).

PAYNE-GAPOSCHKIN, C. *Introduction to Astronomy*. Hodder & Stoughton, 1956.

RUDAUX, L., and DE VAUCOULEURS, G. *Encyclopedia of Astronomy*. Batchworth Press, 1959.

RICHARDSON, R. S. and SKILLING, W. T. *A Brief Text in Astronomy*. Chapman & Hall, 1955.

SPENCER JONES, SIR H. *General Astronomy*. Edward Arnold, 1951.

STRUVE, O. *Elementary Astronomy*. Oxford University Press, 1959.

THE SUN

ELLISON, M.A. *The Sun and its Influence*. Routledge & Kegan Paul, 1955.

BAXTER W. M., *The Sun and the Amateur Astronomer*. Lutterworth Press, 1963

MENZEL, D. H. *Our Sun*. Harvard University Press, 1959.

NEWTON, H. W. *The Face of the Sun*. Penguin Books, 1959.

SEVERNY, A. *The Sun*. Lawrence & Wishart, 1959.

THE MOON

BALDWIN, R. B. *The Face of the Moon*. University of Chicago Press, 1949.

FIELDER, G. *Structure of the Moon's Surface*. Pergamon, 1961.

FIRSOFF, V. A. *Surface of the Moon*. Hutchinson, 1960.

WILKINS, H. P. *Our Moon*. Muller, 1954.

WILKINS, H. P., and MOORE, PATRICK. *The Moon*. Faber & Faber, 1961.

THE PLANETS

(General)

KUIPER, G. P. (editor). *The Atmospheres of the Earth and Planets.* University of Chicago Press, 1951.

MOORE, PATRICK. *The Planets.* Eyre & Spottiswoode, 1962.

SPENCER JONES, SIR H. *Life on Other Worlds.* English University Press, 1955.

UREY, H. C. *The Planets.* Oxford University Press, 1952.

(Mars)

DE VAUCOULEURS, G. *Physics of the Planet Mars.* Faber & Faber, 1954.

MOORE, PATRICK. *Guide to Mars.* Muller, 1958.

STRUGHOLD, H. *The Green and Red Planet.* Sidgwick & Jackson, 1954.

(Venus)

MOORE, PATRICK. *The Planet Venus.* Faber & Faber (3rd Ed., 1960).

AURORÆ

HARANG, L. *The Auroræ.* Chapman & Hall, 1952.

STRØMER, C. *The Polar Aurora.* Oxford (Clarendon Press), 1955.

COMETS AND METEORS

FEDYNSKY, V. *Meteors.* Foreign Languages Publishing House, 1957.

LOVELL, A. C. B. *Meteor Astronomy.* Oxford University Press, 1954.

PORTER, J. G. *Comets and Meteor Streams.* Chapman & Hall, 1952.

ECLIPSES

MITCHELL, S. A. *Eclipses of the Sun.* Columbia University Press, 1951.

RADIO ASTRONOMY

DAVIES, R. D., and PALMER, H. P. *Radio Studies of the Universe.* Routledge & Kegan Paul, 1959.

HANBURY BROWN, A., and LOVELL, A. C. B. *The Exploration of Space by Radio.* Chapman & Hall, 1958.

HEYWOOD, J. (Editor). *Radio Telescopes.* B. A. A. Memoir, 1962.

LOVELL, A. C. B., and CLEGG, J. A. *Radio Astronomy.* Chapman & Hall, 1952.

MENZEL, D. H. (editor). *Radio Noise Spectrum.* Harvard University Press, 1960.

PAWSEY, J. L., and BRACEWELL, R. N. *Radio Astronomy.* Oxford (Clarendon Press), 1955.

PFEIFFER, J. *The Changing Universe.* Gollancz, 1956.

PIDDINGTON, J. H. *Radio Astronomy.* Hutchinson, 1961.

SMITH, F. GRAHAM. *Radio Astronomy.* Penguin Books, 1960.

VAN DE HULTZ (Editor). *Radio Astronomy, I. A. U. Symposium No. 4.* Cambridge University Press.

STELLAR ASTRONOMY

(General)

HYNSK, J. A. *Astrophysics.* McGraw-Hill, 1951.

MCCREA, W. H. *Physics of the Sun and Stars.* Hutchinson, 1950.

PAYNE-GAPOSCHKIN, C. *Stars in the Making.* Eyre & Spottiswoode, 1953.

(Galactic Studies)

BOK, B. F., and P. J. *The Milky Way.* Harvard Books on Astronomy. Blakiston Press, 1946.

SHAPLEY, H. *Galaxies.* Harvard Books on Astronomy. Blakiston Press, 1961.

COSMOLOGY, ETC.

BONDI, H. *Cosmology.* Cambridge University Press, 1952.

COUDERC, P. *Expansion of the Universe.* Faber & Faber, 1950.

RONAN, COLIN A. *Changing Views of the Universe.* Eyre & Spottiswoode, 1961.

STAR ATLASES, ETC.

BEET, E. A. *Ball's Popular Guide to the Heavens.* (Revision). Geo. Philip & Son, 1952.

NORTON, A. P. *Norton's Star Atlas.* Gall & Inglis, 1959.

HISTORICAL

ARMITAGE, A. *The World of Copernicus.* Mentor Books, 1956.

DE VAUCOULEURS, G. *Discovery of the Universe.* Faber & Faber, 1957.

DREYER, J. E. *History of Astronomy from Thales to Kepler.* Dover Publications, 1952.

SHAPLEY, H., and HOWARTH, H. E. *A Source Book of Astronomy.* McGraw Hill, 1960.

SPACE FLIGHT

BATES, D. R. (editor). *Space Research and Exploration.* Eyre & Spottiswoode, 1957.

CLARKE, A. C. *Exploration of Space.* Temple Press, 1951

CLEATOR, P. E. *Into Space.* Allen & Unwin, 1954.

"WIRELESS WORLD" PUBLICATIONS

AMOS, S. W., B.SC. *Principles of Transistor Circuits.* 2nd Edition.

CAMIES, B. S. *Principles of Frequency Modulators.*

COLEBROOK, F. M., B.SC. *Basic Mathematics for Radio and Electronics.*

SCROGGIE, M.G., B.SC. *Principles of Semiconductors.*

SCROGGIE, M. G., B.SC. *Radio and Electronic Laboratory Handbook.*

Radio Valve Data. 7th Edition.

SOWERBY. *Foundations of Wireless.*

LANGFORD-JONES. *Radio Designers Handbook.*

AERIALS & PROPAGATION

KRAUS, J. D. *Antennæ.* McGraw Hill.

Transmission and Propagation. Serving Test Books, H.M.S.O.

WILLIAMS, H. P. *Antenna Theory and Design.* Pitman.

AMERICAN RADIO RELAY LEAGUE. *Antenna Handbook.*

GLOSSARY

ASTEROIDS. Small planets less than 500 miles in diameter, moving in orbits between those of Mars and Jupiter.

ASTRONOMICAL UNIT. A unit of length used by astronomers, being the distance between the Earth and Sun – about $1{\cdot}5 \times 10^{13}$ cm.

BEAMWIDTH. The angle within which an aerial receives or transmits radio waves. The directional diagram or polar diagram is a contour map of the sensitivity of an aerial throughout its beamwidth.

BLACK BODY. A body which absorbs all incident electromagnetic radiation; such a body is also a perfect radiator, the energy radiated at any wavelength depending only on its temperature, according to Planck's Law.

CHROMOSPHERE. *See* Sun.

COMET. A member of the planetary system which exhibits a luminous tail when near the Sun.

CO-ORDINATES. Celestial co-ordinates (Right Ascension and Declination) are fixed among the stars but rotate around the Earth's axis. Galactic co-ordinates are also fixed amongst the stars, but are based on the plane of the Galaxy.

COSMIC RAYS AND COSMIC-RAY ELECTRONS. Elementary particles moving at very nearly the speed of light; cosmic-ray electrons are correspondingly high-speed electrons.

COSMOLOGY. The study of the large-scale structure and evolution of the universe.

db. Decibel. The ratio of input power to output power.

DEGREES KELVIN OR ABSOLUTE (°K). A temperature scale with intervals equal to the Centigrade scale and zero at $-273°$ C.

DIPOLE. The dual active (receiving or transmitting) elements of an aerial.

DOPPLER SHIFT. The frequency shift due to relative motion of a radiating object and an observer. If the distance between the two is decreasing then the frequency is seen to increase; if the distance is increasing, as for external galaxies, then the frequency decreases and the light becomes more red (the "Red Shift").

DRIFT CURVES. Records of extra-terrestrial radiation produced as the beam of an aerial fixed to the Earth sweeps across the sky.

ECLIPTIC. The apparent yearly path of the Sun among the stars.

ELECTRON. The fundamental particle which carries an elementary negative charge and has a mass 1/1836 of the mass of the proton.

FEEDER. The electrical connection between two parts of a radio receiving or transmitting system: for example between the aerial and receiver.

FLUX DENSITY. The strength of a radio (or any other electromagnetic) wave is defined as the amount of power incident per unit area.

GALAXY. The Galaxy is the aggregate of stars (including the Sun), gas and dust, which is visible as the Milky Way; countless other galaxies lie beyond our own system.

GALACTIC LATITUDE AND LONGITUDE. A system of directional co-ordinates in which the galactic plane (the centre of the Milky Way) is zero latitude.

HI, HII. HI is neutral hydrogen and HII is ionized hydrogen (electrons and protons); HI clouds and HIII clouds are interstellar clouds of these gases.

HYDROGEN LINE. The only spectral line (narrow frequency range) so far detected in extra-terrestrial sources in the radio domain. It is emitted by neutral hydrogen at a frequency of 1420 mc/s.

I.C. The index catalogue of nebulæ and star clusters compiled by Dreyer in 1895 (see also N.G.C.).

INTERFEROMETER (Radio astronomy). An arrangement of two aerials spaced a specified number of wavelengths apart and their outputs combined in a receiver.

IONS. Elementary electrically charged particles; most numerous are protons with a positive charge and electrons with an equal negative charge.

IONIZED GAS. A gas in which the atoms have been so excited that each has lost at least one electron.

IONOSPHERE. The part of the Earth's atmosphere above about 50 miles; so-called because some atoms are broken into ions by the solar ultra-violet and x-radiation.

ISOPHOTE. A line on a contour map joining points which represent regions of equal brightness.

LASER. Light-amplification by stimulated emission of radiation.

LIGHT YEAR. A distance unit defined as the distance traversed by a ray of light in one year – $9{\cdot}5 \times 10^{17}$ cm.

LOBES. A fringe pattern created by the use of two aerials combined in an interferometer.

MC/S. Megacycles per second. A description of a frequency of a million cycles per second. The associated wavelength is given by the relation.

$$\text{wavelength in metres} = \frac{300}{\text{frequency in mc/s}}$$

METEOR. A small particle a member of the solar system which enters the Earth's atmosphere at a high supersonic speed. Air friction heats it until it vaporises and it becomes visible as a shooting star.

MASER. Micro-amplification by stimulated emission of radiation.

NEBULA. An astronomical object other than a single star or a cluster of stars; it may be an external galaxy, or a cloud of gas or dust.

N.G.C. The New General Catalogue of nebulous objects prepared by J. L. E. Dreyer in 1888.

NOISE. The effects produced by spontaneous and random electrical fluctuations in radio receivers. The term is used whenever such fluctuations are important, even when they are not converted into audible sound.

NOVA. "New" star whose brightness is increased more than a hundred thousand times during a great explosive outburst.

ORBIT. The path of a heavenly body whose motion is controlled by gravitational attraction of some other body or bodies.

PHOTOSPHERE. *See* Sun.

POLAR DIAGRAM. *See* Beamwidth.

POLARISATION. There is a direction of oscillation of light or radio waves which is always at right angles to the direction in which the wave is moving. *Linear Polarisation*: the direction is constant. *Circular Polarisation*: The direction rotates as the wave advances. In unpolarised radiation the direction is random.

PROTON. A fundamental particle carrying an elementary positive charge whose mass is $1 \cdot 67 \times 10^{-24}$ gm.

PLASMA. A gas whose atoms (some or all) are ionized.

QUANTUM. The smallest unit of energy which can take part in any physical process. The amount of energy is proportional to its frequency.

RADIATION. Energy moving with the speed of light and associated with a band of wavelengths.

RADIO NOISE. Any body which is not absolutely cold generates radio waves or oscillations of a random nature and spreads over a wide band of frequencies. If these waves are amplified and made audible they provide a hissing sound and so are called radio noise.

RED SHIFT. *See* "Doppler Shift".

SOLAR MASS. It is convenient to measure the masses of stars and nebulae in terms of the mass of the Sun – about 2×10^{33} gm.

SPECTROSCOPE, SPECTROGRAPH. Devices for observing and recording spectra, the relative strength of radiation (light or radio) at different wavelengths.

SPECTRUM. The range of wavelengths which make up a beam of radiation.

STERADIAN. The unit solid angle being $\frac{1}{4}$ the solid angle subtend by the whole surface of a sphere at its centre.

SUN. The central star of the solar system. Its visible surface is called the photosphere. Just above this is the cooler chromosphere and above this the very rarefied and hot corona.

SUPERHETERODYNE. A type of radio receiver in which the frequency of the received signal is changed before amplification.

SUPERNOVA. A stellar explosion a thousand times more powerful than a Nova, which increases the brightness of a star ten million times and which blows off a substantial fraction of the star's mass.

SYNCHROTRON RADIATION. The continuum radiation which is emitted by high-speed free electrons or protons when a magnetic field forces them to follow a spiral path.

WAVELENGTH. The distance between two adjacent crests of a wave motion.

YAGI. A special aerial consisting of a driven element and parasitic directors and reflectors.

ZODIACAL LIGHT. A faint glow among the stars in the constellations of the Zodiac caused by sunlight reflected by interplanetary dust and electrons.

INDEX

INDEX

RADIO ASTRONOMY FOR AMATEURS

Amateur Astronomer's Library, Volume II

edited by Patrick Moore

CONTENTS

also containing

76 line diagrams

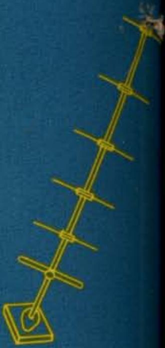
13 photographs

Tables, Charts, Glossary and Index

Published by

LUTTERWORTH PRESS

RADIO
ASTRONOM
FOR
AMATEURS



FRANK
W. HYDE

LUTTERWORT