Chapter 3 - Instruments and Methods of Radio Astronomy

The basic instrument used in radio astronomy, the radio telescope, is a directive aerial system to which a sensitive receiver or receivers are connected: the excellence of its performance is determined primarily as with optical telescopes, by the size of the aperture. Ideally it should:-

(i) Collect a sufficiently large sample of the radiation to make possible the detection of very faint sources. The gain of an aerial system is given by $\frac{4\pi A}{\lambda^2}$, where A is the effective collecting area at wavelength λ. The area should therefore be as large as possible, and if it is a reflecting surface the departures from the correct shape should be less than 1/8th of the shortest wavelength at which it is to be used.

(ii) Have sufficient resolution to permit individual sources, or parts of an extended source, to be selected separately for observation.

The angle, θ, between two sources which can just be separated by a telescope is given by λ/D radians, or $\frac{3438\lambda}{D}$ minutes of arc, where D is the width of the aperture. This angle is usually referred to as the beamwidth, and is approximately the width of the main lobe of the response pattern between the half-power points, as illustrated in Fig. 1. As an example, a parabolic reflecting dish of aperture 10 metres has a beam width of just over 4° when used at a wavelength of 10 cms. The resolution requirements therefore indicate that the aerial should be as large as possible.

(iii) Be capable of being used over a wide range of wavelengths without major modification (other than changing the receiver).

This feature is desirable for determining the spectra of sources, and for search ing for and observing emission at frequencies whose exact value is not known - e.g. because of Doppler shifts, or previously undetected radio frequency lines.

(iv) Be able to be pointed at will over a large fraction of the sky, and to track, or follow, a celestial source with precision.

This is essential for determining the characteristic of sources whose output fluctuates, and to achieve maximum sensitivity by allowing the receiver output to be integrated for extended periods.

(v) Be suitable for determining the polarization of the incoming radio waves.

The form of radio telescope which is a close analogue of the optical telescope is the filled-aperture device in which all the energy falling upon it from a particular (sharply defined) direction is brought to a common focus and compounded to form an image. The simplest and most common form of such a device is the paraboloidal reflector or "dish", the essential feature of the parabolic shape being that parallel rays falling upon it are brought to a single point focus, independent of wavelength. The paraboloidal reflector possesses most of the desirable features listed above. Hertz was the first to use a reflecting dish, in his early experiments with electromagnetic waves, and the first to use such a reflector in radio astronomy was Reber. Paraboloids originally designed for radar purposes were common in the early days, with diameters ranging from 6 to 20 feet, but sizes have gradually increased over the years, the most recent and largest being that of 333 ft. at the Max Planck Institute at the University of Bonn, West Germany. However, the most important aspects of the performance of a radio telescope are determined by its aperture measured in terms of the wavelength at which it is to be used, and the relatively long wavelengths used in radio astronomy inevitably set a limit to the size of the reflecting dish which it is practical or economic to construct. There is thus a limit to the resolution - and also the gain - that it is feasible to obtain with a single filled-aperture device. Even today no radio telescope has a resolution which is substantially better than that of the naked eye at optical wavelengths - about 1 minute of arc. The need for beteer performance, particuarly higher resolution, at the longer wavelengths has been met by devices which are known as interferometers, because their operation is based on the interference between signals which reach the receiver by different paths.

Interferometers

An interferometer in radio astronomy is a device which makes use of two or more separate but interconnected aerials in such a way as to achieve a resolution equivalent to that of a complete or filled aperture of which they can be considered as part. Figure 2 illustrated diagrammatically a simple interferometer in which two aerials separated by a distance D are connected to a receiver midway between them. Signals arriving from a direction at right angles to the base line will reach the receiver together and reinforce each other. Those arriving at an angle θ will only arrive at the receiver in step if the difference in path length, d sin θ, is an integral number of wavelengths. Likewise, at angles θ where d sin θ is an odd number of half wavelengths the signals will reach the receiver exactly out of step, and will thus cancel each other out. The response diagram of such a device is therefore a multi-fingered lobe pattern in which the angle between separate maxima is given by D sin θ = ph where p=1, 2, 3 etc. For small departures from the line at right angles to the base line, the separation between lobes is $\frac{\lambda}{D}$, or if the spacing between aerials is n wavelengths, then the beam width is $\frac{1}{n}$ radians.

Figure 2 shows the response from the receiver as a point source moves through the lobe pattern, and also that from a source which is large enough to be resolved, i.e. shows an appreciable disk. In the latter case when part of the source is in the central part other parts will be in the adjacent regions of zero response, so the amplitudes of the maxima are reduced and there will be no zeros between them. The series of maxima and minima in the receiver response are referred to as fringes, and the ratio of the amplitudes at minimum and maximum, known as the fringe visibility, provides information about the angular size of the source. If the source is a large uniformly bright area maxima from some parts of it will coincide with minima from other parts and the fringe pattern will be smoothed out and may disappear.

The width of the central lobe of an interferometer is a small praction (NM) of that for the component aerials used individually (A/n where the aerials are n wavelengths apart). As an example, a dish of aperture 10 metres has a beam width of approximately 8.5° when used at a wavelength of 1.5 metres, but for two such dishes used as an interferometer with a spacing of 1500 metres (1000 wavelengths) the beamwidth is reduced to 3.4 minutes of arc. The position of sources can thus be determined much more precisely by means of an interferometer. A major difficulty, however, is to decide when the source lies in the main (central) lobe, and this difficulty is considerably accentuated when several sources are present in the region of sky being observed. Some of the early catalogues of radio sources derived from surveys carried out with interferometers contained many erroneous entries from this cause. Nevertheless a number of ingenious devices based on interference principles have been developed which provide outstanding resolution - almost invariably however at the expense of reduced performance in other respects.

The first occasion on which interference principles were used in radioastronomy was in the experiments of Pawsey and his associated near Sydney in 1947 when radio waves from the rising sun reached their cliff top aerial by two paths, one direct and the other after reflection from the surface of the sea. This is illustrated in Figure 3, from which it can be seen that the latter signals are equivalent to those which would be received by a second aerial as a mirror image of that at the cliff top. The device thus acts as a two-aerial interferometer of spacing twice the height of the cliff.

Australian radio astronomers have made a number of outstanding contributions to the design of radio telescopes. One of the first of these was the cross-shaped antenna system developed by B.Y. Mills at the C.S.I.R.O. Radiophysics Laboratory, Sydney, in 1952, to provide an approximate pencil beam for survey work at metre wavelengths; it has become universally known as the Mills Cross. It consists of two long arrays of dipole aerials in the form of a cross, as shown in Fig. 4. These are connected to a receiver through a switch which joins them alternately in phase and out of phase. The outputs are compounded and the net result is that the only signals recorded are those from the small area where the beams overlap. The device thus produces an effective "pencil beam", which can be "steered" in declination by varying the connections to the dipoles in the N-S arm. The principle of the Mills Cross has been followed in a number of other major installations, including a larger version of the Cross by Mills himself and the University of Sydney at Hoskinstown, near Canberra (see page and Plate). [See Images 1-3]

W.N. Christiansen, also of the Radiophysics Laboratory, Sydney, used the principle of a diffraction grating in his "grating interferometer (1955) to provide a means of scanning the Sun and so determining the brightness variations across its disk. His original instrument used 32 small dishes in line but a later version known as the crossed grating interferometer effectively combined the principle of the grating with that of the Mills Cross. It consisted of 64 paraboloids of diameter 19 feet arrayed in the form of a symmetrical cross, the 32 aerials in each arm being equally spaced along a line of length 1200 feet. Each arm produces a series of knife-edge beams which intersect and phase switching is used, as in the Mills Cross, so that only the signals from the intersecting "pencil beam" areas are recorded. The spacing of the aerials is arrnged so that these beams are 1° apart, i.e. only one of them falls on the Sun at a time. As the earth rotates this beam passes across the Sun, scanning a strip in the process: the system of beams is then adjusted so that an adjacent strip is scanned; and so on until the Sun's disk has been covered (see Fig. 5). This was the first radio telescope to produce a television-type radio picture of the Sun - although the picture rate was slow, taking some 40 minutes. The principle of the Crossed Grating Interferometer has been copied in a number of other installations, notably by Christiansen himself in a device intended for cosmic studies which has been erected near Sydney by the Department of Electrical Engineering of the University of Sydney (see page and Plate). [See Image 4]

Radio telescope systems in which spaced aerials are combined to simulate the resolution of a filled-aperture device of very large aperture appear to offer the only means of obtaining extremely narrow beams, the limit being set - at least for the time being - by the maximum separations possible on the earth. Unfilled aperture devices can be divided into two broad types according to whether they produce an image point-by-point (i.e. by image synthesis) or by compounding the Fourier components derived by observing at a series of different spacings (aperture synthesis). The chief characteristics of filled and dilute aperture aerials will now be described, with illustrations of typical examples.

Filled-Aperture Instruments

(i) The first of these devices designed especially for radio-astronomy was that built by Grote Reber in the backyard of his home in Wheaton, Illinois, in 1937 (see Plate I, page : it was a steerable parabolic reflector, 31 ft. in diameter and focussed the radiation on to dipole at its focus. Aerials of this type are simple electrically and can be used at any wavelength merely by changing the focal feed and associated receiver. They point at the source being received, so the quality of the image and the resolution is independent of position in the sky; and their circular symmetry makes them suitable for polarization measurements. Parabolic dishes have therefore been widely used in radio astronomy. In recent years a number of successful dishes up to 300 ft. in diameter and of high surface accuracy have been built. The cost however increases approximately as the weight, i.e. as the cube of the linear dimensions, and the difficulties of maintaining correct shape as the telescope is tilted also increase with size. Paraboloidal dishes of diameter substantially larger than 300-400 feet in diameter, though mechanically feasible if servo-control of the shape of the reflecting surface were incorporated, are therefore unlikely to be built unless some means becomes available of overcoming the problems associated with the law of weight proportional to the cube of the diameter.

Some of the outstanding radio astronomy installations using large reflecting dishes are as follows:

Jodrell Bank. The first of the large steerable dishes was the 250 ft instrument of the University of Manchester at Jodrell Bank, built in 1956, and originally designed for metre-wave observations. Some modifications were introduced following the discovery of the hydrogen line, but the dish is of limited use below 21 cms. Its distinguishing feature is a solid metal reflecting surface of very short focal length. A MKII reflector of much higher surface accuracy was completed at Jodrell Bank in : this is elliptical in shape   and good for use down to   cms.

Parkes, Australia.   One of the most successful of the large dishes is the 210 ft paraboloid of the Australian National Radio Astronomy Observatory at Parkes, N.S.W.: it was commissioned in 1961 and has been shown to perform with good efficiency to at least 6 cms wavelength.   A novel feature of the design is use for the first time of servo control devices to maintain pointing of the dish accurately in line with that of a precision equatorally-driven axis.   This has proved so successful that a number of large dishes constructed since that time (e.g. two telescopes of diam 210 ft built by the U.S. National Aeronautics and Space Administration) have incorporated this master equatorial control system.

Greenbank, West Virginia.   For many years the largest parabolic dish in existence was the 300 ft instrument at the American National Radio Astronomy Observatory at Greenbank, completed late in 1962.   This is a transit instrument, i.e. fixed in the north south meridian, scanning in right ascension being provided by the earth's daily rotation.

A high-precision dish of diameter 140 ft was brought into operation at Greenbank in   : this has been used with good efficiency at wavelengths as short as   [See Images 5-7]

Bonn, West Germany.   What will be for some years to come at least, the largest steerable radio telescope anywhere in the world is a 100 metre (333 ft) dish in the sparsely populated Eifel Mountains, south west of Bonn, West Germany. It is operated by the Astronomical Institute of Bonn University and is due to be completed in 1970.   The surface accuracy in the central region 60 metres in diameter is suitable for operation at a wavelength of 3 cms and efficient operation over the full aperture is possible down to 5 cms.   The Whole structure turns in azimuth on a single circular track, as for the Jodrell Bank Mark I instrument .It is illustrated in Plate [see Images 8-9]

Non-Steerable Reflectors

One approach to reducing the cost of large reflecting telescopes has been to use a fixed reflecting surface which can be supported from the ground at many points, rotation of the earth being used to provide for observing at different right ascensions. The largest of these devices is in Puerto Rico where a vertical-pointing spherical reflector 1000 ft in diameter has been constructed in a natural bowl-shaped hollow in the ground, suspended above the surface by relatively short supports (see Plate VIII).[image 10] The spherical surface allows the beam to be moved up to 20° from the zenith by moving the feed, but the latter needs to be a very complicated device to remove the spherical aberration which is inherent with a reflector of this shape.

Another device in operation at the University of Illinois is built in a natural gully, with the reflecting mesh placed directly on the ground. This is in the form of a parabolic cylinder with dimensions of 600 ft by 400 ft (equivalent to a circular dish of diameter 122 metres) and intended as a survey instrument for use at a fixed frequency near 600 MHz. The beam can be swung electrically between declinations 10°N and 70°N, the earth's rotation providing sweeping in right ascension.

Another approach to filled apertures is that first introduced by J.D. Kraus at the University of Ohio in 1953. This makes use of a large flat reflecting surface which can be tilted so as to direct the radio waves from the desired source on to a vertically-mounted sector of a parabolic mirror which has a collecting device at its focus. The width of the reflector is much less than its height so the beam is fan-shaped. A later version of this type is in operation at Nancay, France,[image 11] and is suitable for operation down to about 10 cms wavelength. The fixed reflector is 115 ft high and 1000 ft long. Another form of this type is at the Pulkovo Observatory, and uses only one reflecting surface, which is made up of a large number of small flat sheets which can be positioned to form a strip of a parabolic surface. This antenna also has a fan-shaped beam and can be used down to wavelengths as short as 3 cms.

Other Types

Arrays of dipoles are satisfactory for image formation but are unsuitable when high directivity is desired because of the extreme complications involved in the electrical problems associated with connecting the millions of individual dipoles that would be required. Horns of large dimensions are very expensive to construct, but their directivity can be calculated with some accuracy and their use is almost entirely confined to reference standards or calibrators.

(ii) Unfilled Apertures.- Image synthesis.

The first unfilled aperture device to be used in radio astronomy was the cliff interferometer used by McCready, Pawsey and Payne-Scott with which the association between solar radio frequency radiation and sunspots was discovered at Sydney in February 1946. It consisted of a radar array of effective area 9.5 metres^2 located on a cliff 278 feet above sea level, together with its image in the sea below. The equivalent aperture was thus 556 feet. In announcing their discovery the same authors pointed out that it is possible, in principle, to derive the brightness distribution, i.e. an image, of a source by synthesis of the Fourrier components produced by varying the spacing between the two antennas. The first to actually do this for the Sun was Starvier in 1950.

Mills Cross.

The first unfilled aperture radio telescope to produce effectively a pencil beam was the original Mills Cross, developed by B.Y. Mills at the Australian Radiophysics

Laboratory in 1952. This operated at a wavelength of 3½ metres and consisted of two strip antennas, each 1500 feet long, arranged as a horizontal cross: each arm had a fan-shaped response and when the output of both was multiplied the resultant response was that of the part of the sky common to the two beams. A later version of this device - the Molonglo Cross - has been constructed by Mills at Hoskinstown near Canberra, with arms in the form of cylindrical paraboloids 1 mile long (see Plate IX). It is designed to operate at wavelengths of 73.5 cms. and 2 metres and has a resolution of 3 minutes of arc at 73.5 cms. It is fitted with a number of receivers and when used for transit observations can record up to 11 points simultaneously.

Other examples of the Mills Cross type have been built at Bologna, Italy, and Serpukhov in the U.S.S.R. The former operates at wavelengths and with arms 1 km long and has a resolution of minutes of arc: the latter has a resolution of 3minutes of arc at metre wavelengths.

<u>The Culgoora Radioheliograph</u>

A novel form of unfilled aperture has been developed by J.P. Wild of the Australian Radiophysics Laboratory especially for detailed studies of the Sun. This consists of 96 separate steerable parabolic aerials arranged around the perimeter of a circle of diameter 1.9 miles and has a resolution of 4.3 minutes of arc at its operating wavelength of 3.75 metres (80 MHz) see Plate X. Its special feature is that, under the control of an inbuilt computer, the aerials are connected to 48 receivers (with appropriate phases) so that 48 beams are produced simultaneously along a north-south line. This line of points is then made to sweep rapidly in an east-west direction, taking up 60 different positions in each second. The radioheliograph thus produces a two dimensional picture of the Sun derived from the raster of 48x60 = 2880 points recorded each second. This facility of allowing rapidly-varying phenomena on the Sun to be observed has made possible major advances in our understanding

Transcribed in 2022 by Miller Goss to improve legibility

Page 11 of Chapter 3 “Instruments and Methods of Radio Astronomy” by A J. Higgs, draft of never published book *Radio Astronomy*, September 1968

….Laboratory in 1952. This operated at a wavelength of 3 ½ metres as a horizontal cross, each arm had a fan-shaped response and when the output of both was multiplied the resultant response was that of the part of the sky common to the two beams. A later version of this device- the Molonglo Cross- has been constructed by Mills at Hoskinstown near Canberra, with arms in the form of cylindrical paraboloids one mile long (see plat IX). It is designed to operate at 73.5 cm and 2 metre and has a resolution of 3 minutes of arc at 73.5 cm. It is fitted with a number of receivers and when used for transit observations can record up to 11 points simultaneously.

Other examples of the Mills Cross type have been built at Bologna, Italy and Puschino in the USSR. The former operates at wavelength 73.5 cm and with arms 564 metre long and has a resolution of 4 minutes of arc; the later has a resolution of about 3 minutes of arc at metre wavelengths.

The Culgoora Radioheliograph: A novel form of unfilled aperture has been developed by J.P. Wild of the Australian Radiophysics Laboratory especially for detailed studies of the sun. This consists of 96 separate steerable parabolic aerials around the periphery of a circle of diameter 1.9 miles and has a resolution of 4.3 minutes of arc at its operating wavelength of 3.75 metres (80 MHz). (see plate X) Its special feature is that, under the control of an intelligent computer, the aerials are connected to 48 receivers (with appropriate phases) so that 48 beams are produced simultaneously along a north-south line. This line of points is then made to steer rapidly in an east-west direction, taking up 60 different positions in each second. The radioheliograph thus produces a two-dimensional picture of the sun derived from the raster of 48x 60=2880 points recorded each second. This facility of allowing rapidly-varying phenomena on the sun to be observed has made possible major advances in our
understanding….

of complex processes in operation in the solar atmosphere (see Chapter).

(iii) Unfilled apertures - Aperture Synthesis

Another variation of the spaced-aerial technique which makes possible the construction at reasonable cost of radio telescopes of even higher resolution than those described above is that in which a fixed aerial is combined with one which is movable and can take up a large number of other spacings and orientations. The movable aerial thus occupies different parts of the full aperture at different times and the complete image is built up by sYnthesis of the necessary Fourier components. The name "aperture synthesis" was first given to an antenna system of this kind developed by Ryle and Hewish. To simulate the resolution obtainable with a filled aperture device of diameter D the movable aerial must traverse in sequence a semicircular area of radius D centred on the fixed aerial, the patterns of amplitude and phase being measured and recorded at each position for compounding with those derived at all other positions. With this device almost unlimited resolution can be attained with D sufficiently large. A severe limitation, however, is that synthesis of the full image is likely to take an inordinately long time, even though in practice the movable aerial need not take up every possible position within its range.

The Cambridge Aperture Synthesis Telescope

A number of interferometer instruments fall in the aperture synthesis category but the historic antenna of this kind is that constructed at Cambridge, England, by Ryle and Hewish in . It consisted of a fixed cylindrical paraboloid 1450 ft x 65 ft and a smaller aerial 190 x 65 ft which was movable on railway tracks at right angles to the fixed unit. A later version of the aperture synthesis instrument being used by Ryle and his colleagues at Cambridge uses three antennas in line separated by 2500 feet. Each is a 65 foot diameter paraboloid, two being fixed and the third movable on rails. The rotation of the earth is used to sweep the baseline through 180° in 12 hours: about 50 different spacings on 50 different days are needed to complete the image of a region of sky about 1° across. The resolution is equivalent to a filled aperture of elliptical shape whose major axis is 5000 ft, i.e. 29 seconds of arc at 21 cms wavelength in the optimum direction [See image 12-13]

The Christiansen Compound Interferometer

By the addition of four fully steerable paraboloid dishes of diameter 45 feet to his original crossed grating interferometer of 64 19 ft paraboloids described on page Christiansen has produced a device capable of producing a map of a one-square-degree area of sky in a shorter time (approximately one day) than is possible with any other aperture synthesis instrument. The outputs from each arm of the original cross together with that from the 45-ft dishes are fed to an electronic computer where they are appropriately compounded to produce the information from which the final contour map is drawn. Plate XI is a view of the aerial system. [See Image 14]

The Westerbork Array [See image 15]

A project which began as the "Benelux Cross" and has subsequently been carried through entirely by the Netherlands is due to come into operation in 1970. This employs a total of 12 paraboloids of diameter 83 ft: ten are fixed along a base line of length 5000 feet and the other two are movable on tracks at right angles to the base line. The large collecting area of the 12 dishes will give this radio telescope the highest sensitivity of any of the existing installations.

Summary

The steerable paraboloidal reflector yields high quality images which are independent of the size and complexity of the sources being observed is completely flexible; and is suitable for all types of radioastronomy observation, the only qualification being that consideration of cost and mechanical feasibility set an upper limit to the aperture and hence the gain and resolution that can be attained. Larger effective areas are possible with fixed reflectors and hence higher resolution but with loss of mechanical and (usually) electrical flexibility, arrays of large numbers of reflecting dishes can be used to provide fine pencil beams and lend themselves to multibeam operation and hence rapid scanning: there is no mechanical limit to the apertures possible, but the electronic complexities increase rapidly at the number of separate reflectors involved.

The unfilled aperture arrays are attractive when the prime requirement is high resolution for surveying weak sources of small angular dimension. The highest resolution possible can be provided by means of aperture synthesis provided an immediate answer is not required. The disadvantage of these devices is the very low rate of observation, determined by the time required to synthesise all the necessary Fourier components over an extended observing period.

Receivers

The radiation collected by even a large aerial system is almost always too weak to be detected at the focus and so a receiver is required to provide sufficient amplification to raise the signals to a level which will operate the recording device being used. The power level from distant sources is extremely low and often well below the "flux unit" used by radioastronomers (defined as 10^{-26} watts m^{-2} $(c/s)^{-1}$) so considerable amplification is usually required. There is little difficulty in achieving this with the electronic devices now available, and without loss of essential information about phase and spectrum contained in the incoming signals. A wide variety of radio receivers is used in radioastronomy, most of them operating at a designed wavelength, or over a limited timing range. A basic difficulty is that the incoming wanted radio signals carry no distinctive modulation since they arise from essentially random processes at their source, and so are indistinguishable in quality from the unwanted noise signals generated by the random motion of the electrons in the components of the receiver. The unwanted noise signals arising within the receiver are amplified equally with the cosmic signals, and set a limit to the amplification that can be used, and to the sensitivity of the receiver, i.e. to the smallest signals that can be distinguished. Great advances in receiver design that have occurred during the past decade have come from reductions in internal receiver noise, and have made possible the attainment of higher sensitivities at some wavelengths that can be availed of. A practical limit to the sensitivity is reached when all the sources which the aerial is capable of resolving can be detected: when this point is reached no more sources can be worked on no matter how much more sensitive the receiver may be.

Noise Fluctuations

When a radio telescope is pointed at a source the output of its associated receiver will contain contributions not only from the source being observed, but from fluctuation noise arising within the receiver itself and the associated feed and transmission line from the sky background, including other sources too faint to be resolved and also from the ground. The thermal noise generated per unit band width from the fluctuation noise within a resistance at temperature T ^{o}K can be shown to be K.T (where K is Boltzmann's constant) and it is usual to specify the signal components in terms of the temperature of a resistance at the aerial input terminals which would give the same output: the aerial temperature due to the source is usually referred to as T_a and the remaining components as a systems noise temperature T_N. The chief component of the latter is almost always that arising within the receiver. The merit of a receiver is indicated by its effective noise temperature which should obviously be as low as possible: many of the early discoveries in radio astronomy were made with receivers having noise temperatures of 5000°K or so, in marked contrast with figures as low as 3.5°K which are currently attainable. An alternative method of specifying receiver performance is in terms of a <u>noise factor</u>, which is defined as the ratio of the actual power output from the receiver to that which be given by an ideal noiseless receiver (in which the only output would be that from the fluctuation noise in its aerial).

The "wanted" signals from the source are indicated by an increase in the total power output from the receiver when the source is in the beam. This level is constantly flucuating however, in particular because of the random character of the various components contributive to the total output, and there is a limit to the minimum signal that can be distinguished, determined by the amplitude of the fluctuations about their mean level. Some of the fluctuations may be due to variations in the overall amplification of the receiver and a great deal of care is taken to minimize these by stabilizing the power supplies and operating critical components at constant temperature. An effective way to reduce them still further is to employ a comparison system (such as that due to Dicke) in which the receiver input is switched rapidly (e.g. at 400 times per second) between the aerial and a dummy resistance or other calibrating noise source maintained at nearly the same temperature as the aerial, so that only the difference signal is recorded. The flucauations which remain are due almost entirely to the random nature of the collisions between electrons which are responsible for the original signals.

If the receiver accepts a band of frequencies B cycles per second then the detector output consists of approximately B independent contributions per second. The fluctuations can be substantially reduced by averaging or integrating them (by adding a condenser across the output) over suitable periods, which can vary from seconds to minutes as long as the wanted signals are not of a transient nature. By this means the residual root-mean-square fluctuations can be reduced by a factor $\frac{1}{\sqrt{Br}}$ where r is the total time over which averaging takes place.

As an example, with a receiver of bandwidth 20 Mc/s which has a noise temperature of 150° and used with an aerial system where the sky background is 2° and the ground spillover 10° (i.e. a total systems noise temperature of 162°) the r.m.s. fluctuations can be reduced, by integration over 5 seconds, by a factor of $\frac{1}{\sqrt{20 \times 10^6 \times 5}}$ i.e. 1000. The smallest signal that can be reliably detected in the presence of noise is usually taken to be about three times the r.m.s. fluctuations, i.e. with the receiver used as above this would be about .05 °K. It may seem remarkable that signals can be detected when their level is such a small fraction of that of the ambient noise, but this is a commonplace in radio astronomy.

Types of Receiver

Radioastronomy began in the metre and decimetre wavebands, using sensitive receivers developed for radar use during World War II, and gradually spread to shorter wavelengths as the necessary techniques and components were developed. The need for extremely sensitive receivers at these wavelengths has indeed provided a realistic stimulus to designers - and this has led to remarkable advances which, of course, will ultimately benefit the electronic industry as a whole. A variety of techniques is needed to cover the range for 3 mm to 300 metres.

For observations in the continuum the need to reduce noise fluctuations calls for wide bandwidths and figures of tens to hundreds of megahertz can profitably be used, associated with time constants that provide integration times of 10 seconds or more. Receivers for spectral line observations, on the other hand, need to accept signals over a very much narrower band whose width is determined by the spectral resolution required. Spectral lines usually show Doppler shifts corresponding to the velocity of the source in the line of sight, and to delineate this in the case of some OH clouds, for example, requires bandwidths as narrow as 1 kHz. In the best receivers elaborate precautions are taken to ensure the highest possible stability within the receivers and, in addition, comparison techniques are employed. In continuum receivers the receiver is switched rapidly between the aerial and a reference noise source and only the difference signal recorded; while in spectral line receivers the timing alternates rapidly between two frequencies, one within the line profile and the other in the continuum just outside the line.

Superheterodyne receivers developed for radar were used in the early radioastronomy observations, often with crystal detectors. The need to operate at higher frequencies where the cosmic noise level becomes progressively lower was met initially by the development of low-noise triodes and later of semi-conductor devices, with crystal mixers still popular as detectors at the shortest wavelengths because of their relative simplicity and good stability.

Improvement in low-noise gun-structures for travelling-wave tubes made possible good sensitivity at centimetre wavelengths because of the extremely wide bandwidths available with travelling wave tubes. These tubes have not become popular, however, because of the difficulty in obtaining adequate mechanical stability in their internal structure and hence of reducing gain fluctuations to an adequately low level to make use of the wide bandwidth available. A major step forward came with the introduction of devices which exploit inherent low-noise properties rather than bandwidth. These are based on the use of physical principles not previously applied to the design of amplifiers, and draw their power from a radio-frequency source called the ump" and usually higher in frequency than the signal to be amplified, rather than from a D.C. source as in conventional amplifiers. There are two distinct types, the "maser" (abbreviated from Microwave Amplification by the Stimulated Emission of Radiation) and the "parametric amplifier": both make use of the special properties available in solid-state devices.

The Maser

The operation of the maser can be explained in terms of the quantum theory, in accordance with which the energy of an atom is restricted to a number of discrete levels. When it changes from one level to the next higher energy level a quantum of energy is absorbed; if it moves to a lower energy level radiation is emitted. Radiation quanta of appropriate value falling on a suitable substance may be absorbed by some atoms and raise them to a higher level, or it may cause atoms already at the higher level to emit radiation, which will be in phase with the original radiation and so increase it or result in amplification. This is called "stimulated emission". Under normal circumstances there are more atoms in lower energy states than in the upper, but in a maser the action of the "pump" is to provide quanta which will increase the population of the upper levels so that when the wanted signals pass through the material some of the atoms in the higher levels will spontaneously emit quanta of radiation and so increase the energy of the wanted signals, i.e. produce amplification.

A number of materials in the form of single crystals have energy levels which can be used in this way, one of the most suitable being ruby. When maintained in a steady magnetic field of about 2000 gauss this material provides amplification at Hydrogen line frequencies around 21 cms. One of the great disadvantages of the maser however is that it is necessary to operate at very low temperatures if sueful amplification is to be obtained. This is because the noise temperature of the maser and the bandwidth characteristics are better when the degree of population of the higher levels is greatest, which occurs at liquid helium temperatures. The whole device is usually immersed in a dewar flask containing liquid helium and the problems involved in arranging for the necessary facilities have limited the application of masers, particularly since parametric devices are capable of comparable performance. Maser receivers have been constructed with a noise temperature of , and with a travelling wave version down to , which is approaching the ul timate performance at microwave frequencies.

Parametric Amplifiers

The parametric amplifier is not dependent upon quantum theory: it derives its name from the fact that rhythmic variation of one of the receiver parameters (usually a capacitance) is used to produce amplification. The action is analogous to that of a child who is building up his motion on a swing: he raises his centre of gravity by straightening his legs each time the swing passes through its rest position (thus increasing the amplitude of that oscillation) and lowers his centre of gravity when he reaches the end of that swing. The energy acquired by the increased amplitude of swing is provided by the boy himself and is applied twice per cycle, i.e. at double the frequency of the swing's oscillations.

In a parametric amplifier the swing is replaced by an electrical circuit consisting of a capacitance C and inductance L: connected in parallel the condenser is alternately charged from and discharges into the inductance, at a frequency LC. If the capacitance is suddenly reduced when it is fully charged this will increase the voltage across it, thus increasing the energy in the system. If the capacitance is increased to its former value when the condenser is discharged, i.e. when there is no voltage across it, no change in total energy of the system occurs. The energy is the system can thus be increased by varying the capacitance at double the frequency to which the circuit is tuned. In practice the resonant circuit is tuned to the frequency of the signals to be amplified and the capacitance varied electronically in one of a number of ways, the most effective being by the use of semiconductor diodes, called "varactors" in which the desired variations are a function of the bias on the diode.

Parametric amplifiers avoid many of the difficulties found with masers and are widely used in radioastronomy: in the decimetre range they provide a performance comparable with that of maser systems at a fraction of the cost and which at longer wavelengths meets most of the requirements. At room temperatures noise temperatures below 50 $^{\circ}K$ can be achieved, while cooling with liquid nitrogen has reduced noise in the amplifier to below 20 $^{\circ}K$ and in liquid helium to still lower values. In most cases however the noise contributed by other receiver components essential for stability, together with sky background and ground radiation, will exceed that generated within the amplifier itself. Representative values for a good parametric amplifier for use at 21 cms, coded in liquid nitrogen are:- amplifier noise temperature 25° and ground spillover 30° and remaining systems noises 50°.

Tunnel Diodes

The tunnel diode is a semi-conductor device which has the attractive feature of developing a negative resistance characteristic merely by the application of a low power D.C. bias voltage. It is thus suitable for use in negative-resistance amplifying circuits without the need for the pump-supply required for masers and parametric amplifiers. Its noise temperature at microwave frequencies however is relatively high ($\sim 250^{\circ}$) but the simple circuiting involved makes it attractive, particularly for use when the sky background temperature is high. It is also useful as a mixer where, in contrast with a conventional crystal in this application, it can give appreciable conversion gain.

Correlation Techniques

Correlation techniques have found increasing application as an essential part of some of the large and costly arrays now in use, particularly in association with suitable digital computers. In the Australian Mills Cross for example the outputs from the 177 frequency converters and IF preamplifiers in the N-S array, and the 22 mixers and preamplifiers in the E-W array, are combined with appropriate phase delays and multiplied so as to produce 11 simultaneous beams at different declinations. By correlating a number of phase arrangements of the outputs for the E-W array with those for the N-S array the resulting series of beams can be used to form a simultaneous "image" of a region of the sky. This technique is used extensively in aperture synthesis devices, e.g. as developed at Cambridge, England.

Correlation techniques have also been used in spectral line receivers e.g. as developed by Weiwieb at the U.S. National Radio Astronomy Observatory at Greenbank. The receiver noise output is auto-correlated, the necessary delays being derived from separate receivers, delay lines or digital computers. The resultant output is the spectrum of the receiver noise plus that of the radiation being received.

Precise Location of Radio Sources

The demand for precise position fixing of radio sources came with the realization that there are other sources in the skies besides the Sun, and that these are not associated with any of the bright stars. Progress in elucidating the nature of the sources and the processes which result in the emission of radio waves so much more strongly than light waves was obviously dependent upon identifying some at least of them which could also be studied by optical techniques. The first such identifications with the remains of a supernova and with two somewhat unusual galaxies were promising enough, but a major stimulus came two years later when the strongest source in the skies Cygnus A was found to be coincident with a strange object so distant as to appear optically the 18° magnitude, i.e. only visible with the aid of the largest telescopes.

An essential requirement for accurate location of radio sources is that /the searchaing device should have high resolution, i.e. a very narrow "pencil" beam, preferably circular in section arther than fan-shaped, and receive energy from no other directions, i.e. have negligible side lobes. Associated with this should be enough sensitivity to bring a reasonable number of fainter sources within range of detection. Both these requirements are met from filled-aperture devices, particularly large steerable parabolic dishes but, as we have seen above, practical considerations limit the performance that can be achieved to some 5 minutes or arc at 10 cm wavelengths, and 40 minutes at 1 metre. Some of the best survey and identification work has been done with the Parkes 210 ft dish, in the course of which a grid of calibrator sources has been produced from those of essentially point dimensions whose identification with optical objects is regarded as certain. The position of the calibrators is thus known with high precision, and they serve as absolute reference points against which the position of other sources can be determined with comparable accuracy.

For higher resolution at longer wavelengths, it is generally necessary to use interferometer devices. In addition to considerable electronic complexities inherent in the more sophisticated examples of this type of radio telescope are the uncertainties of using a device whose beam is not circularly symmetrical and can have more than one prominent lobe. Great care is therefore needed in specifying positions as derived with these instruments: some of the early surveys included a number of errors from this cause.

A technique which does not suffer from any of these defects and is yet capable of the highest precision is one which is based on the occultation of sources by astronomical bodies whose position is known with high accuracy. This method conveniently provided by nature has been confined almost exclusively to the Moon and was used with dramatic success by Hazard and Mackey with the 210 ft Parkes Telescope to observe the quasar 3C273, in 1962. The observation consists of noting the times of immersion and emersion and recording the signal strengths. If this be done the Moon's position is known well enough to enable that of the source to be specified to within seconds of arc. In addition, the Moon also acts as a diffracting edge and variations in signal strength as the source enters and emerges from eclipse can provide information about source structure and the size of the components (see Figure 5). By this means the mysterious quasar 3C273 was first located with sufficient precision to enable it to be identified, and shown to have a composite structure which matched that of its optical counterpart (see Chapter).

Angular Size of Sources

The early identification of a few of the discrete sources with nebulae but none with visible stars suggested that the majority, perhaps, were not of stellar dimensions. The first radio evidence of this came in 1951 during a survey by Mills in Sydney at a wavelength of 3 metres: he found that several sources gave a smaller response on a widely-spaced interferometer than on a narrow one, indicating an extent of about a degree for these sources. Piddington and Minnett, also in Sydney, discovered about the same time the existence of another source in Cygnus which was about 2 degrees by 7 degrees in extent - they called it Cygnus X to distinguish it from Bolton's Cygnus A source. Independent evidence as to the size of several of the brighter radio sources came in 1952 from observers in Cambridge and Manchester. Since then interferometer measurements at gradually increasing base lines have succeeded in putting dimensions on a large proportion of the known radio sources.

By 1962, however, a small number of sources had been found for which the corresponding object on a star chart for that area appeared to be of stellar dimensions. These objects therefore came to be known as quasi-stellar objects and are discussed in more detail in Chapter . The problem of accounting for their very great luminosity has led to an intensive effort to determine their angular size. A number of these objects showed no sign of being resolved by interferometer systems of increasingly great separations, in which the component aerials were linked by land-line and radio links, and it has been necessary to separate them by intercontinental distances: this is known as "very long base line interferometry", or V.L.B.I., and has only been made possible by the availability of atomic clocks (particularly using caesium and rubidium) which are capable of maintaining a sufficiently constant rate to serve as time-marking devices at each end of the chosen base line. The outputs from very stable receivers at each station which have been tuned to exactly the same frequency (i.e. after allowing for the differing Doppler shifts at the two stations due to the earth's rotation) are recorded on magnetic tapes, together with timing marks from their atomic clocks, and the two tapes are subsequently synchronised (with the aid of the timing marks) and compounded. Fringes are observed, as with a more conventional interferometer, if the source dimensions are less than the resolution corresponding to the station spacing.

The precautions necessary for success in V.L.B.I. work are enormously more complex than indicated in this simple outline but this technique has already been used successfully een Parkes, Australia, and California and Greenbank in the nd awa, Canada (see Fig. 6), i.e. over baselines sent the maximum length that can be obtainable on

Final paragraph of Higgs page 25, transcribed to improve legibility by Miller Goss, 2022

The precautions necessary for success in V.L.B.I work are enormously more complex than indicated in this simple outline but this technique has already been used successfully between Parkes, Australia and California and Greenbank in the US and Ottawa, Canada (see Fig 6) i.e. over baselines comparable to the maximum length that can be obtained on earth.

[See Images 16-17 for images of OVRO, a major component of the US VLBI Network.]