Correlator Arrays — A New Generation Of....

RADIO TELESCOPES

by G. W. Swenson, Jr.

Departments of

Electrical Engineering

and Astronomy

University of Illinois at Urbana-Champaign

Although astronomy is probably the oldest of the sciences, it is presently in a dynamic state of developement, thanks to a number of recent technological advances in instrumentation. Among the most fruitful of these developments has been the radio telescope in its various forms, which in thirty-five years has extended the observable spectrum of cosmic electromagnetic radiation from five octaves to sixteen octaves. It has made visible certain parts of the universe which are obscured by interstellar dust clouds at optical wavelengths, and has added several new observable physical quantities to the astronomer's inventory.

The two principal purposes of a telescope, whether for optical or for radio astronomy, are to intercept an easily detectable amount of electromagnetic energy from the cosmic source under study, and to produce an image of the source sufficiently magnified to yield the desired amount of detail. Unfortunately, neither radio nor optical telescopes have ever been able to meet these criteria well enough to satisfy completely the requirements of astronomical research. In both radio and optical cases, compensation for limitations on energy-collecting ability resulting from limited area can be accomplished to a certain degree by extended observing time.

The limitation on the detail that can be distinguished, that is, the angular resolution of the instrument, is a more fundamental limitation. In the optical case the limit is imposed by the inhomogeneous atmosphere of the earth, and under the best "seeing" conditions is about one-half second of arc. The Rayleigh criterion for angular resolution states that two point sources are resolvable if they are separated by at least half the angle between the first-order nulls of the telescope's Fraunhofer diffraction pattern. The angular resolution, in radians, is equal to the reciprocal of the diameter in wavelengths of the objective lens (or mirror) of the telescope. For a typical optical wavelength, a telescope whose objective aperture is 20 centimeters in diameter has half-second resolution in the absence of atmospheric effects. Thus, the reason for building telescopes up to 4.5 meters in diameter is to improve light-gathering ability rather than angular resolution. Telescopes on space vehicles outside the earth's atmosphere, or on the moon, can now yield much better resolution, and sophisticated methods of processing optical data can overcome, in part, the effects of poor "seeing". For ground-based observations it is important to select a site at a high altitude which is relatively free from the worst of the atmospheric inhomogeneities which limit angular resolution.

Radio telescopes are, for the most part, not limited in angular resolution by atmospheric effects. Rather, the limits to date have been imposed by economic and engineering considerations. Radio telescopes typically utilize wavelengths a million times as great as those of optical telescopes; hence, for comparable performance, the physical dimensions must also be a million times as great. To achieve a resolution of one-half second at a wavelength of 10 cm requires an aperture diameter of approximately 42 km. Clearly it is not possible to con-

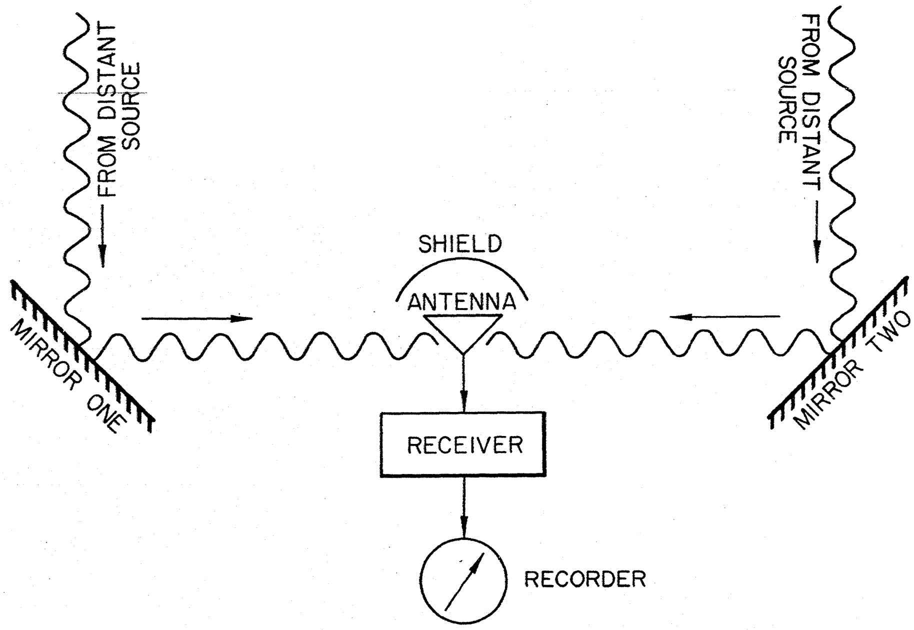

Fig. 1 An Elementary Radio-Interferometer

struct a lens or a mirror of this size, or even of one-tenth this size, and it is thus unknown from direct observational evidence whether or not atmospheric effects would impose a resolution limit on an antenna of this type.

The best resolution achieved to date with a single-reflector radio telescope, in the range of wavelengths above 5 mm, is about one minute of arc, achieved by a few telescopes in the wavelength range 0.8 to 3 cm.

In order to overcome the structural and monetary limitations of the size and resolution of single-reflector instruments, radio astronomers have, almost since the birth of their science, utilized a general principle by means of which great angular resolution can be obtained at relatively small cost. This technique, called radio interferometry, is analogous to that of the Michelson interferometer of optics.

One form of such an instrument is shown schematically in Figure 1. Two mirrors, well-separated along an East-West line on the ground, reflect waves from a cosmic source into an antenna, which is itself shielded from waves arriving directly from the source. As the source is at such a great distance from the telescope, the incoming waves may be assumed to follow parallel paths. It will also be assumed that the distance from each mirror to the receiver is exactly the same. Both mirrors rotate about their centers in such a way that the waves from the source are always reflected to the antenna. If the source is on the meridian, that is, in the north-south vertical plane, both waves require exactly the same length of time to travel from the source to the antenna. If, however, the source is east of the meridian the wave arriving via the west mirror takes somewhat longer than does the other component. At the antenna, the electric field is the sum of the two incoming components. If the source is at the meridian, the two components are in phase and the electric field at the antenna is twice that of free space. When the difference in length between the two paths is an odd number of half wavelengths, the two components at the antenna are of opposite phase and add to zero. The result of the westward motion of the source is a steady oscillation in field strength at the antenna, so that the output of the receiver is also sinusoidal in time.

The foregoing description is based upon the assumption of a source of infinitesimal diameter, emitting radio waves on only one frequency (or wavelength). Cosmic sources (galaxies, planets, nebulae, the sun) generally emit incoherent radiation with broad power spectra, and, or course, have non-zero angular dimensions. The antennas and radio receivers used in radio astronomy respond only to relatively narrow bands of wavelengths; thus, the theory of the radio-astronomical interferometer deals with spatially incoherent, band-limited (quasi-monochromatic), electromagnetic waves. Researchers in the theory of incoherent waves have shown that the response of an interferometer to such radiations is a sinusoidal oscillation of the output of the receiver, similar in many respects to the response of the simplified system we have described.

In order to understand the application of the interferometer to a more general problem, let us imagine a source consisting of two points of equal strength, separated by a small angle. First, let the angular separation be very small, so that the interferometer output due to one source is nearly in phase with the output due to the other source. Then, assuming a linear system, the indicated output will be nearly twice the output from a single source. Next, let the angular separation be increased until the outputs due to the two points are substantially out of phase. The phase difference is (approximately) proportional to the angle between the points, so long as this angle is small. The resultant output for this case is substantially less than twice the output due to either point source alone. Finally, let the angular separation between source components be such that output phase difference is 180°. Then the resultant is zero. The interferometer is, in a sense, "blind" to this particular source.

Clearly, the response of the interferometer depends upon the size and shape of the source. For the simple, two-component source assumed here, the separation of the source components can be determined from the output response of the interferometer. The angular separation to which the instrument is "blind" is directly related to the difference between the two path lengths, and this, in turn, depends upon the separation, in wavelengths, of the two mirrors. By moving the two mirrors, one can find the separation that produces the "blind" response, and hence the angular separation of the two sources.

A few additional words of jargon will be useful in our further discussion. The sinusoidal output traces of an interferometer are usually called "fringes", which are characterized by "amplitudes" and "phases". An interferometer is said to measure the "fringe visibility" of a source, the visibility being, in general, a function of the baseline length; that is, the distance between the mirrors. For a single point source, the visibility is constant in both amplitude and phase and, as we have seen, for a complex source the visibility can be strongly dependent on the baseline length.

It is interesting to consider the case of a number of pairs of mirrors, all reflecting waves from the same cosmic source into the same receiving antenna. In effect, we have a number of interferometers, all utilizing the same receiving system. If all the mirrors are tangent to a parabolic curve (Fig. 2) whose plane contains the cosmic source and whose focus is at the antenna, the path length from the source to the antenna is the same, regardless of which mirror is involved. For a source on the axis of the parabola the electric field at the antenna can be considered to be the superposition of the contributions of the several interferometers. Alternatively, it could be considered that we have constructed a parabolic reflector. However we view the matter, it is clear that the physical result is the same; thus, there must exist a close relationship between the results obtained by scanning a source with a number of interferometers of different baselines, and those obtained by a parabolic reflector of fairly conventional design.

It should be mentioned that it is unnecessary to place the plates tangent to the parabola. Similar results can be obtained by placing them near the ground, properly oriented to reflect the incoming rays into the antenna, and so adjusted in height that the various ray paths from source to antenna differ only by integral numbers of wavelengths. This scheme is seldom used because, for any particular installation, it operates satisfactorily over a rather narrow band of wavelengths. The idea does serve, however, to emphasize the equivalency between an assembly of interferometers and a continuous-aperture antenna.

We have now established that the information gathered from a number of interferometers, in terms of fringe amplitudes and phases, is somehow equivalent to the information gathered by a more conventional reflector-type radio telescope. The information from each interferometer can be expressed completely, for a given source, by the two numbers representing the fringe amplitude and phase. The antenna and receiver somehow combine this information to yield a picture of the source. It should also be possible to record the information obtained by scanning the source with several interferometers, one at a time, and to process the information by hand or in an automatic computer, to yield the same picture of the source. This, in fact, is the technique usually employed in radio interferometry. Clearly, what we are using here is the familiar technique of harmonic analysis, described in 1822 by Joseph Fourier in his classical treatise, Théorie Analytique de la Chaleur. In fact, the interferometer with variable baseline gathers the necessary information about the source to permit us to write a Fourier Series representing the brightness distribution of the source.

The Fourier Series has been the subject of a tremendous amount of mathematical research and is very well understood. Briefly, Dirichlet's theorem states that if a function is periodic in the independent variable and reasonably well-behaved, it can be represented with complete accuracy by an infinite series of harmonically-related sinusoidal terms with appropriately chosen amplitudes and phases. The accompanying figure (Fig. 3) illustrates the principle for a particularly simple periodic function.

On the area scale of interest to us here, the brightness of the sky is not a periodic function of the celestial coordinates. However, if it is possible to disregard the actual brightness of the sky outside an interval of particular interest, we can assume that the function is periodic outside the interval. The green-shaded portion of the figure represents the actual brightness distribution of the source. This particular function can be completely specified by the sum of three terms: a con-

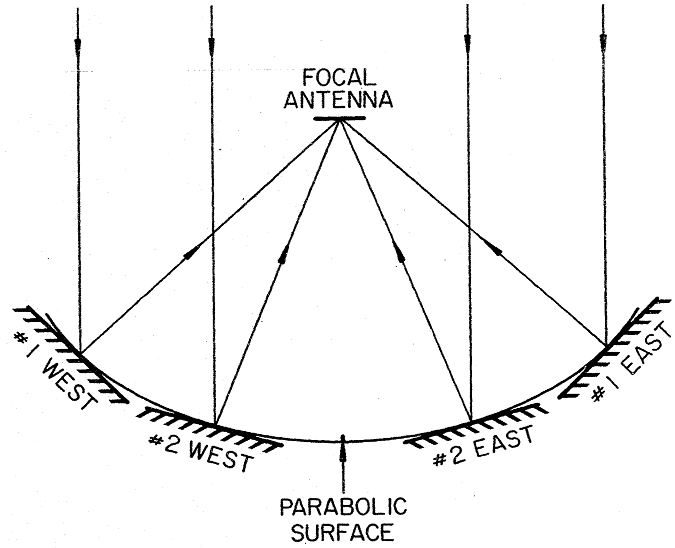

Fig. 2 A Form of Multiple Interferometer

stant term equal to the average value of the function, a "fundamental" term whose spatial period is equal to the interval of the sky to be analyzed, and a "third harmonic" term whose amplitude is 25% of that of the fundamental. A more general type of function would require more terms for an adequate approximation; in the extreme case in which abrupt discontinuities occur (they must be finite) an infinite number of terms is needed. An infinite number of terms would require infinite baselines; the higher the spatial frequency of a term the longer is the corresponding interferometer baseline. The longest baseline, then, determines the finest source detail that can be measured. Up to this point we have assumed that the sky brightness distribution is one-dimensional. In the actual, two-dimensional case, it is necessary to scan the source in two dimensions. This complicates the problem enormously. Fortunately, the existence of an adequate mathematical theory and of large-capacity automatic computers make the technique perfectly feasible.

It is now clear that the maximum angular resolution achievable with a given ensemble of interferometers depends upon the longest available baseline. Further, since the shortest baseline needed is determined by the maximum width of the region of sky to be examined, and since all other baselines are multiples of this length, the number of different baselines needed can now be determined. For a one-dimensional source distribution (to be scanned by a one-dimensional array of interferometers) it is the width of the desired region of sky di-

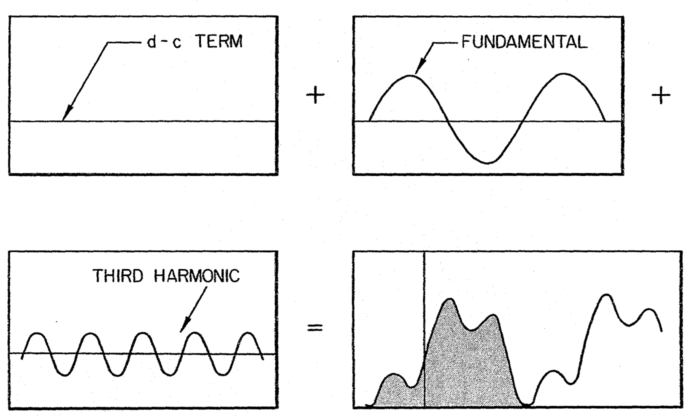

Fig. 3 Fourier Synthesis

vided by the desired angular resolution, each expressed in the same angular units. For example, to synthesize the brightness distribution of a strip of sky 10 minutes of arc wide, to a resolution of 3 seconds of arc (1/20th of a minute), requires a Fourier series of 200 terms, and thus, 200 different interferometer baselines.

In astronomical practice, interferometers consisting of mirrors are seldom used, although the earliest such experiment (McCready, Pawsey, and Payne-Scott, 1947) was conducted on a seaside cliff, with the surface of the sea as a mirror. It is much more satisfactory to replace the mirrors with antennas and to interconnect them with electrical cables, as in Figure 4. The antennas can track the source synchronously, and should be large enough so that their narrow beamwidth excludes the sky outside the small area being synthesized. This latter requirement is a difficult one. In fact, the theory requires that, if all interferometers were present simultaneously, their antennas would touch one another, edge to edge. Fortunately, there is a way of escaping this requirement, and this will be discussed later.

In designing a large array of interferometers it is necessary to have a systematic method of determining which of the Fourier components of the brightness distribution of a source can be supplied by the array, in order to determine whether the proposed instrument is adequate for the measurements intended. To derive such a method it is useful to have a graphical representation in which the spatial frequency of each Fourier component, and the direction on the sky in which it is measured, can be represented as a point. As each Fourier component corresponds uniquely with the length and direction of the interferometer baseline used to measure it, it would also suffice to plot the baselines in our array. Each baseline has a length and a direction, as projected upon the celestial sphere; therefore, it can be represented as a vector. Let us imagine that our array is to observe a source in its zenith, so that radio waves from the source travel to the earth along a path perpendicular to the plane of the array. Then let us plot the various baselines in a cartesian coordinate system, as in Fig. 5, in which each dot represents an interferometer baseline. The distance from the origin is the length of the baseline and the direction from the v-axis is the azimuth of the baseline from the north direction.

Now let us suppose that we desire to map an extended source which is confined to an area of the celestial sphere φ radians in diameter. The lowest spatial frequency needed in the Fourier series is then $1/\Phi$ cycles per radian, and all terms in the series are harmonics of this frequency. This corresponds with a baseline length of $1/\Phi$ wavelengths. Let the desired angular resolution be b radians; then the maximum required baseline length in any direction is 1/b wavelength. Thus, in order to measure all of the Fourier coefficients we need, we must fill a circular area of our u-v diagram (Fig. 5), 2/b wavelengths in diameter, with dots spaced $1/\Phi$ wavelengths apart. A circular array of dots in the u-v diagram yields equal angular resolution in all directions on the celestial sphere. The total number of dots in the circle is approximately $\pi(\Phi/b)^2$. Only half of this number of baselines is actually needed; the symmetry of the u-v diagram results in a two-fold redundancy of points because, while a baseline does have a direction, it does not have a positive or negative sense.

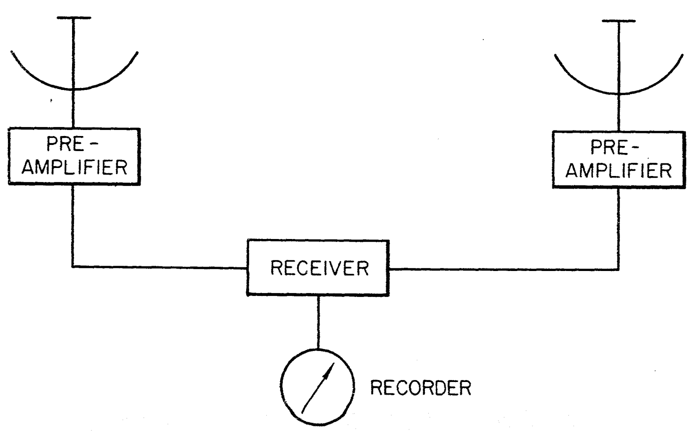

Fig. 4 A More Practical Form of Radio-Interferometer

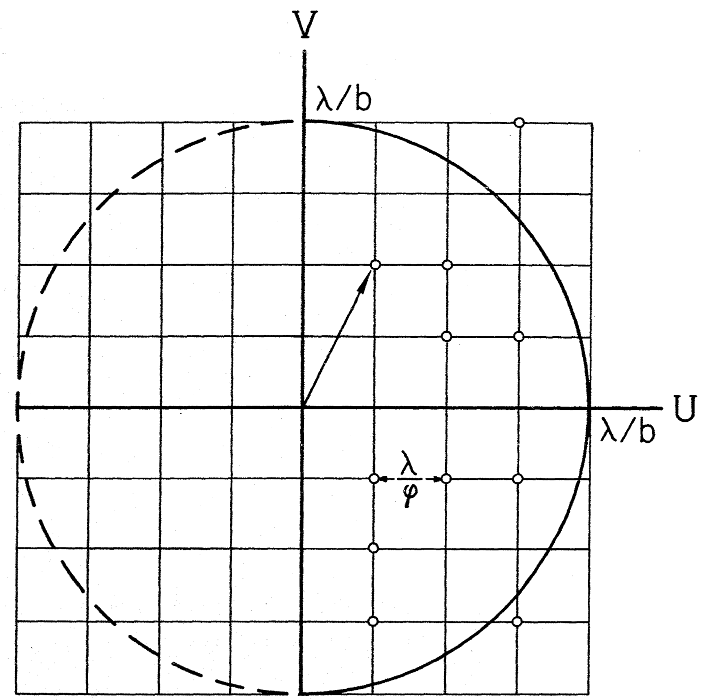

Fig. 5 The U - V Plane

As an example of the use of the u-v diagram, we may evaluate a cross-shaped array consisting of 32 antennas, with a uniform spacing of 200 wavelengths. The overall length of each arm is thus 3000 wavelengths. It is assumed that a correlator is connected between each possible combination of two antennas. Figure 6 shows the resulting u-v diagram. The numbers and letters indicate the existing baselines. The grid is based upon a supposed requirement of a field of view, Φ, of 5 minutes of arc and an angular resolution, b, of 10 seconds of arc. Each square of the grid, falling within the circle of radius 1/b, should contain an entry if we are to measure enough Fourier coefficients to synthesize our chosen piece of sky with the desired amount of detail. As it happens, it is impossible to design an array of many antennas which produces only one entry per square. There is an inevitable problem of redundancy. Clearly the number of empty squares shows that the example array is inadequate for the job, and either the number of antennas must be increased or a number of separate measurements must be made with different arrangements of the antennas.

So far we have discussed only the case in which the source is in the antenna's zenith. Suppose now that the source is somewhere else on the celestial sphere, such that the line from source to center of baseline makes an angle Θ with the baseline. The length of any baseline, as projected on the celestial sphere, is forshortened in proportion to the sine of Θ. Thus a given horizontal baseline could have a projected length anywhere between its actual length for a source in the zenith, and zero, for a source in the direction of the baseline. If a source is observed only at meridian transit, each projected baseline will have only one length, but if all the antennas in the array track the source across the sky as the earth turns, all baselines will vary continuously with time. With a little spherical trigonometry one can show that the diurnal track of the projected-baseline vector is an ellipse in the u-v diagram (Fig. 7). For a source on the celestial equator the ellipse degenerates into an East-West line; for a source at the North or South celestial pole, it is a circle. Thus, each baseline has a number of different values, and every time the elliptical locus passes through one of the squares in the u-v diagram we have acquired another of the necessary Fourier components. Plotting the elliptical tracks for all the baselines in an array shows at once which of the Fourier components is still missing, and the procedure in designing an array is to try a number of possible array configurations, until a configuration is found which yields acceptable results for the desired range of source declinations.

Thus, the use of hour-angle tracking can be used to produce a "synthetic" array whose u-v diagram is essentially completely filled, while requiring far fewer antennas and receivers than would be required for the meridian-transit mode of operation.

In order to increase the "field-of-view" of an array, it is necessary that the length of the shortest baseline be reduced, and, consequently, that a finer grid be used in the u-v diagram. Hour-angle tracking is of great assistance in filling the extra squares resulting from the finer mesh. The computations for plotting the u-v diagram for a large array are complex enough to require the use of a large digital computer which plots the data points directly.

Figure 8 shows the u-v diagram for the cross-array investigated earlier, except that in the present case the individual antennas track the source for eight hours (four hours before meridian transit and four after). The XX, YY, ZZ entries indicate points of high redundancy. Clearly the use of hour-angle tracking provides an enormous improvement in the u-v diagram coverage. To obtain the same performance with the meridian transit mode of operation would require many more antennas and correlators and a correspondingly larger amount of ancillary equipment. Thus, substantial economies in initial costs and operating costs can be achieved by hour-angle tracking, at the expense of greatly increased observing time per completed observation.

The correlator array, consisting of an assembly of interferometers, makes possible a sort of "ultimate" ground-based radio telescope, one whose angular resolution is limited by inhomogeneities in the atmosphere rather than by cost or by the strength of construction materials. One outstanding advantage of the correlator array is that the quality of a single Fourier component can be improved by repeated observations; thus, the angular resolution of the "synthetic" telescope simulated by a large correlator array can conceivably be much better than that of the single-reflector telescope which is equivalent in size. In such a single-dish instrument all Fourier components are observed and superimposed **simultaneously,** and the possibility of improving atmospherically-degraded resolution by repeated observation does not exist.

Several highly-successful correlator arrays have been built to operate in the "source-tracking" or "supersynthesis" mode described above. This technique was first applied to astronomical practice at the Mullard Observatory of the University of Cambridge, England. Other major facilities exist at the National Radio Astronomy Observatory in West Virginia, and at the Owens Valley Radio Observatory of the California Institute of Technology. These three instruments each consist of three or more paraboloidal antennas, movable on linear

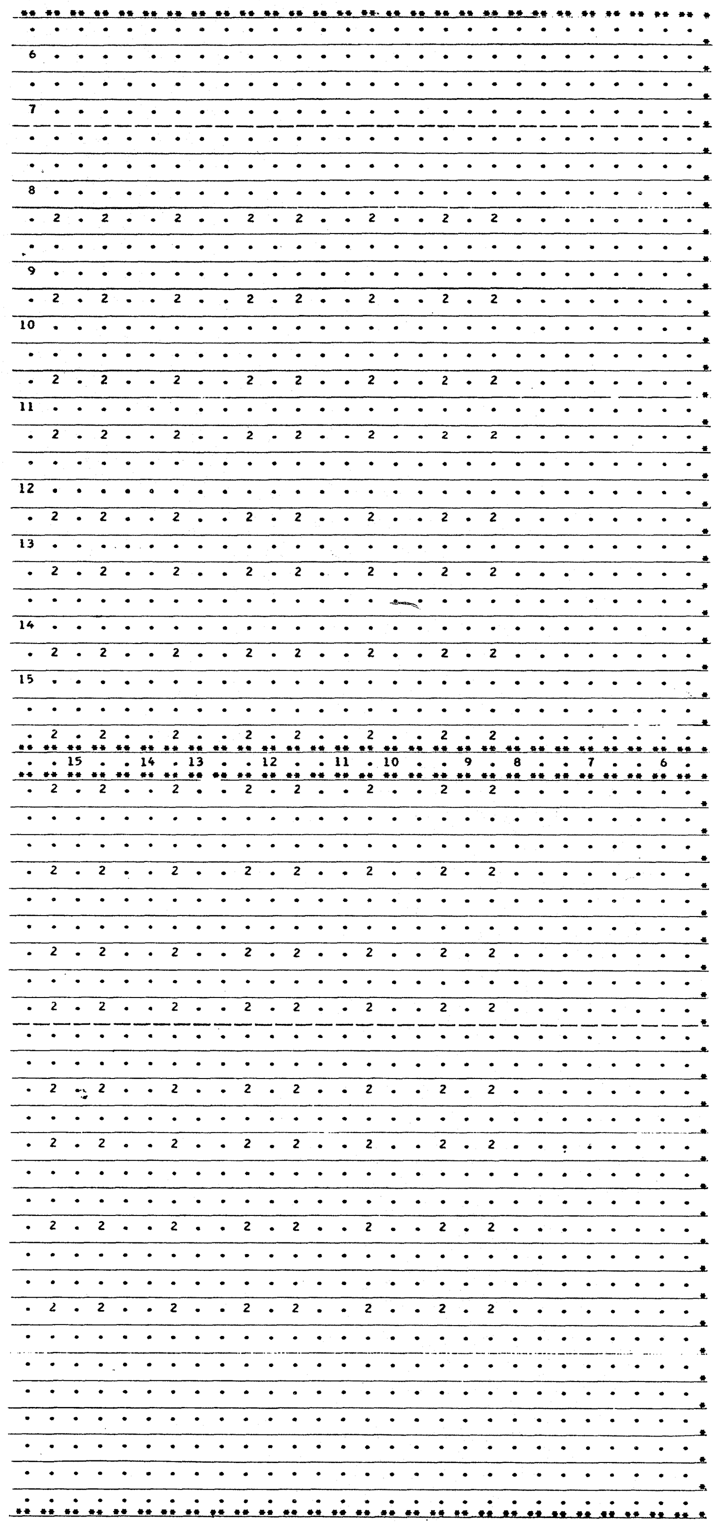

Fig. 6

TRANSFER FUNCTION OF THE CROSS ARRAY - NO TRACKING.

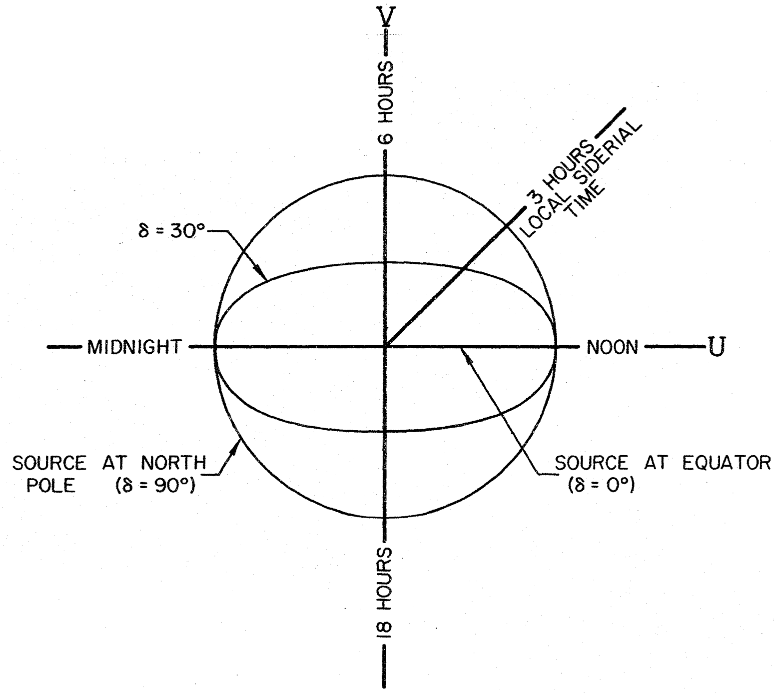

Fig. 7 Locus of Projections on the Celestial Sphere of a Horizontal, East-West Baseline.

baselines out to separations of a few kilometers, giving angular resolutions of a few seconds of arc at 3 cm wavelength. With linear baselines only sources well away from the celestial equator can be successfully mapped.

A five-element linear array of fixed paraboloids has been built at Stanford University, and a cross-shaped array of relatively small dishes at Sydney University in Australia. In the Netherlands a 12-element array of 25-meter-diameter paraboloids has been built on a one-kilometer linear baseline. Two of these antennas are movable. With such a large number of baselines available simultaneously this Dutch telescope is capable of a complete source synthesis in only two days' observing, making it a relatively fast instrument.

The technique of earth-rotation synthesis has thus been very adequately demonstrated at angular resolutions of several seconds of arc. For many years, American radio astronomers have dreamed of an instrument of this kind with a sufficient number of antennas to permit synthesis in one day of a complete picture of a distant, cosmic source. Such an instrument was under study at the National Radio Astronomy Observatory as early as 1964, but no source of funds could be found at that time. The program was finally approved by the U. S. Government in the early 1970's; in the meantime, design improvements and inflation had increased the price to over $70 million. This would be by far the most ambitious and expensive ground-based astronomical instrument ever built. It is by no means certain that the necessary appropriations of money for this instrument will follow in regular installments. If all goes as planned, however, the instrument should be in full operation in the early 1980's, nearly twenty years after its technical form was first conceived, and nearly thirty years after the astronomical need for it was perceived. The instrument is known as the VLA, for Very Large Array. It will occupy a Y-shaped configuration of railroad track, thirteen miles on a leg, in a high mountain basin of New Mexico. It will operate on several wavelengths in the centimeter range, and will produce pictures with detail as small as a quarter of a second of arc. The technical realization of this very sophisticated instrument involves many state-of-the-art techniques in electronics, such as dozens of cooled solid-state parametric amplifiers, circular-mode broadband-waveguide transmission, ultra-high-speed digital information transmission and processing, complete computer control of the entire operation, and extremely elaborate on-line data processing.

While all this has been in the planning process, radio astronomers have been extending the interferometer principle to baselines of continental and intercontinental length, achieving angular resolutions as fine as 10^{-4} second of arc. In 1967, radio astronomers of Canada and the USA demonstrated successful operation of interferometers with baselines hundreds or thousands of miles long, without any electrical connections between their constituent radio telescopes. Signals at each telescope were recorded on broad-band magnetic-tape recorders, together with extremely stable and precise time signals. Later, the pair of tapes was played into a "correlator", a device which multiplies the broad-band signals together after suitably synchronizing them to a fraction of a microsecond. The correlator output is the interference-fringe function of the interferometer. The recording and correlating can be done either digitally or analog-wise.

The initial successes in this new interferometer technique, which is now generally known as VLBI (very-long-baseline interferometry), opened enormous new possibilities for astronomical research. Whereas angular resolution at either optical or radio wavelengths had been limited to a few tenths of a second of arc, VLBI methods immediately made possible resolutions of 10^{-4} second, limited only by the diameter of the earth and the practical short-wavelength limit imposed by atmospheric transparency and telescope precision.

VLBI observations have made possible the investigation of the size and (to a limited degree) the structure of extremely distant sources, such as quasars. The quasars, many of which are stupendous emitters of radio waves, are among the most puzzling objects in the universe. Not only are they extraordinarily energetic, suggesting unknown processes of radiant energy production, but they also include the most distant known objects in the universe. So resistant have they been to the available methods of investigation that it is clear that more detailed mapping procedures are needed.

A two element interferometer can yield only one Fourier component of a source brightness-distribution at a time. If the source can be tracked all day, an elliptical locus is produced, as described earlier, each of whose points represents a different Fourier component. Thus the VLBI can produce some information about the brightness map of the source, and, in fact, can distinguish among certain simple models of a source. Unfortunately, complicated sources do not yield to this approach. In such cases, it is necessary to use more telescopes, in order to produce more baselines and more elliptical loci. The more baselines one has available, the more parameters one can utilize in the model of a source. In principle if one has all the necessary Fourier components, and if both amplitude and phase of each component are known, the source map can be produced by straightforward Fourier transformation. Unfortunately, with VLBI, not only is the number of data severely limited but the phase information is generally lacking completely. This latter misfortune results not only from lack of sufficient precision in the time signals at both ends of the baseline, but also from time variations in the phase delays suffered in the atmosphere by the radio waves reaching the two antennas. Thus, atmospheric effects on continental or global scale potentially limit the resolution obtainable with VLBI techniques.

There are a few possible solutions to the dilemma of the missing phase, none of them completely satisfactory. At any rate, there is no technique other than VLBI which can yield comparable information concerning the most distant cosmic radio sources. Whether a method can be developed for obtaining fringe phase or for synthesizing source maps without using phase, or whether it will be necessary to rely on the present method of testing intuitively-developed source models, it seems clear that relatively-large arrays of antennas operating in the VLBI mode are necessary to progress in understanding quasars and other very distant sources. This understanding, in turn, is necessary to a better understanding of the universe as a whole: its history, constitution, and destiny.

Fortunately, enough antennas of proper type and location already exist in North America to permit the organization of a very respectable VLBI array. The writer has estimated that a capital outlay of $6 million would be sufficient to provide the necessary time-standards, tape recorders, radio receivers and data-processing facilities. The performance of this array, assuming that fringe phase is available, can be described in terms of its synthesized beam pattern, as an alternative to the transfer function used earlier. The (power) beam pattern and the transfer function are Fourier transforms of one another, and the beam pattern can be thought of as the response of the antenna to a signal emanating from a distant point source. The pattern for an array consisting of seven antennas in Maryland, Massachusetts, Southern Ontario, West Virginia, Southern Michigan, Western Texas, and central California is shown in the figure below. Antenna engineers will recognize this as a radial section of the pattern of an antenna with a quite respectable beam shape and sidelobe levels, indicating that the performance is surprisingly good for such a sparsely-occupied array.

Should the public desire to promote research in astronomy in an effective way, the continental array of very-long-baseline interferometers promises extremely large returns in knowledge of the universe at comparatively modest cost. It will surely represent the next generation of radio telescopes, as the smaller synthetic-aperture instruments represent the recently-matured generation.

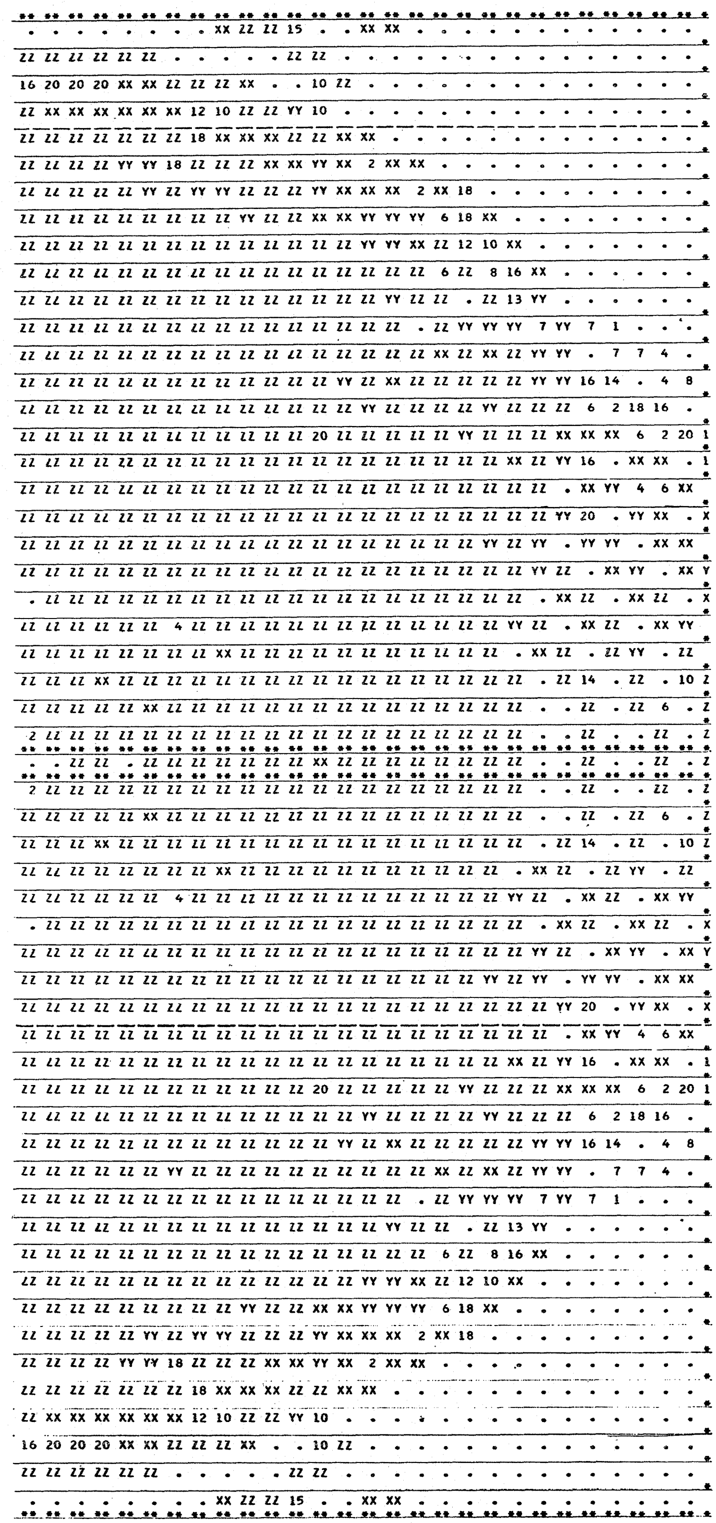

Fig. 8 Transfer Function of the Cross Array — Eight Hours Tracking.

Fig. 9 ➧

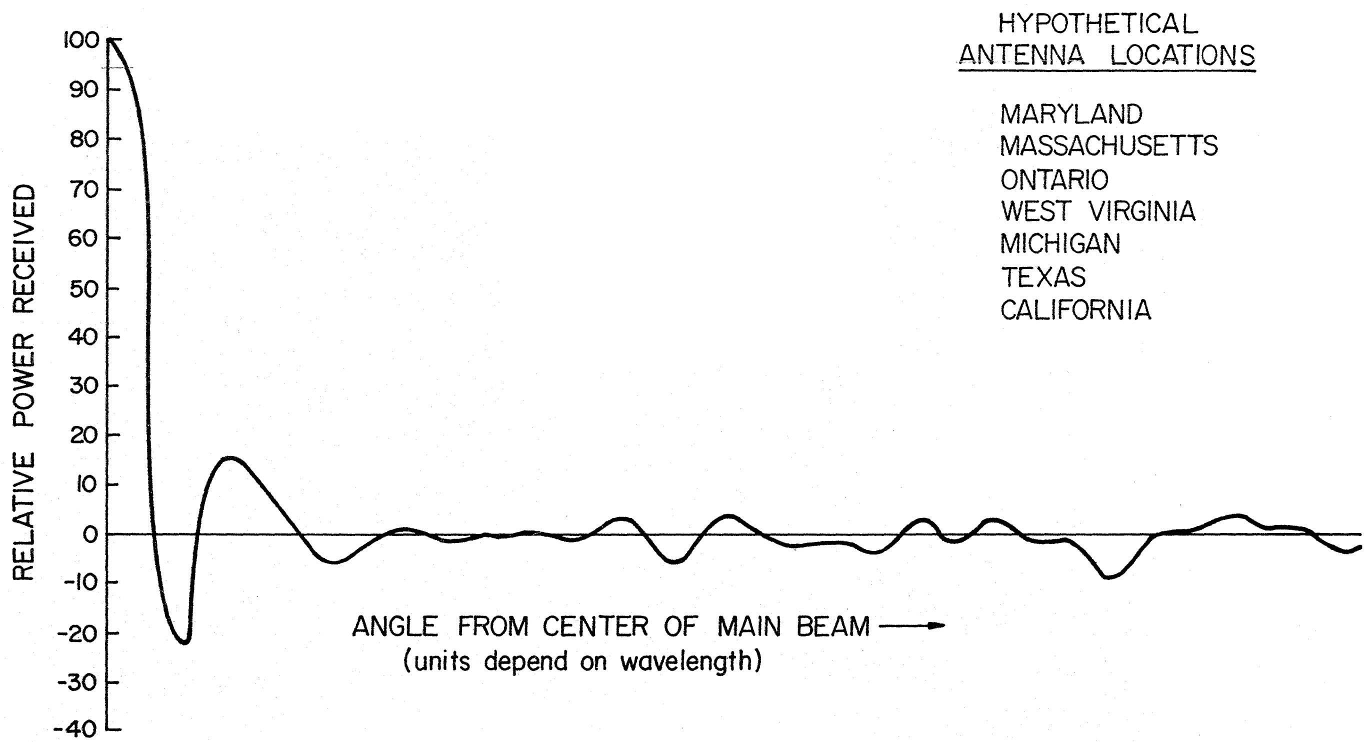

Figure 9. Reception pattern of VLBI Array with six antennas.