535.854 + 523

INTERFEROMETERS IN ASTROPHYSICS

V. I. SLYSH

Usp. Fiz. Nauk 87, 471-489 (November, 1965)

1. INTRODUCTION

THE most important task of observational astronomy is to obtain pictures of celestial bodies. This is a trivial problem when dealing with the sun, planets, and nebulas, for their images are produced on photographic plates by the usual procedure of astronomical observations. This technique, however, cannot yield images of the main objects of optical astronomy, the stars. By "image" we mean a two-dimensional distribution of the density of a photographic plate, corresponding uniquely to the angular distribution of the brightness of the star or of the nebula. While this correspondence is quite good for the sun, planets, and nebulas, the same cannot be said of the stars. What is usually called the image of the star actually constitutes a smeared light spot, having nothing in common with the angular distribution of the star's brightness. This difference is due to the different relation between the angular dimensions of the celestial object and the angular resolution of the instruments used to carry out the observations. The angular dimension of the sun is approximately 30′, while the planets and nebulas measure several dozen seconds, minutes, or degrees. The resolution of modern optical telescopes, on the other hand, is measured in seconds or even fractions of seconds of an angle. The resolution limit of optical telescopes is set by the diffraction of light. For a telescope with diameter $D = 1$ meter at a wavelength 5,000 Å, the diffraction limit is $\lambda/D \approx 0.1''$. Under real conditions, however, this limit is not attained, owing to many limitations imposed by the flicker due to the inhomogeneity of the atmosphere, by imperfections of the emulsion and of the optical systems, etc. By virtue of these causes, modern telescopes are unable to resolve images smaller than 1 second of arc. At the same time, the dimensions of even the nearest stars are much smaller than the resolution of modern telescopes. In fact, a star such as our sun, when seen from a distance of only several parsec, will subtend an angle less than 0.01″. The dimensions of more remote stars are of course even smaller.

Similar difficulties are encountered also in radio astronomy. In the radio band, the resolution is limited primarily by the diffraction, since the wavelength is 5—6 orders of magnitude larger, and the dimensions of the radio telescopes exceed those of optical telescopes by only 1—3 orders of magnitude. The angular dimensions of the objects investigated in radio astronomy range from several or several dozen degrees to several and fractions of a second. These are the sun, the planets, and the nebulas. Optical astronomy and radio astronomy have been recently engaged in studies of celestial objects with stellar angular dimensions, such as the so called quasistellar objects or quasars, and also exploding stars. Theoretical estimates show that some sources of radio emission can have angular dimensions amounting to hundredths or even thousandths of a second of arc [1]. Thus, measurement of the angular dimensions and the production of images of celestial bodies is a most important task in both optical and radio astronomy. The development of astronomy in other regions of the electromagnetic frequency spectrum—infrared, x-rays, and probably also the long-wave region of the spectrum—calls for the development of suitable methods of obtaining images for these regions too. One such method is the observation of the occultation of celestial objects by the moon. Referring the reader to the appropriate article in UFN [2], where its application in the optical region is described, we note only that a successful method has been recently developed for lunar occultations in radio astronomy [3], making possible exact measurement of the coordinates of the radio source 3C-273 and leading to the discovery of the first quasar and to a measurement of the angular dimension and coordinates of the x-ray source in the Crab nebula, thus refuting the hypothesis that this source is a so called neutron star. [4] The method of lunar occultations is limited in scope and therefore cannot compete with the interferometric method.

2. INTERFEROMETERS IN OPTICS AND IN THE RADIO BAND

The idea of using an interferometer in astronomy was proposed by Fizeau back in the middle of the nineteenth century. However, the first direct measurement of the angular diameter of a star was not made until 1920 by Michelson and Pease [5], who measured the angular diameter of Betelgeuse (α Orionis) with the aid of a six-meter interferometer installed on the 100-inch telescope of the Mount Wilson observatory. During the ten years that followed, Pease measured six more stars; all giants or supergiants of later spectral classes, from K0 to M8. Their angular diameters ranged from 0.047″ to 0.020″, calling for a separation from 3 to 7.3 meters between mirrors. Michelson's stellar interferometer consisted of two plane mirrors (Fig. 1), spaced several meters apart, the rays from which were guided into the telescope by

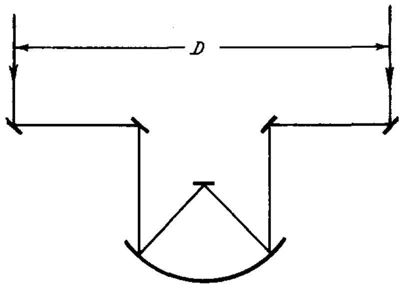

FIG. 1. Michelson stellar interferometer.

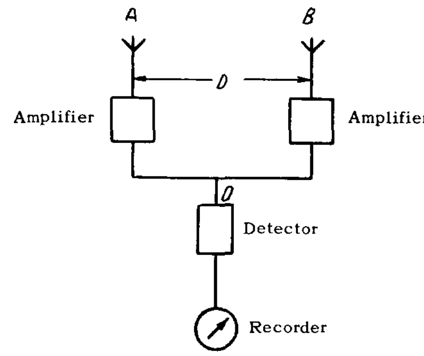

FIG. 2. Michelson radio interferometer.

means of two periscopic mirrors. Interference fringes, whose depth decreased with increasing distance between mirrors as the star was resolved, were recorded on a photographic plate at the focus of the telescope. The resolution of the interferometer was λ/D, where D is the distance between mirrors. Michelson's interferometer is particularly suitable for visual investigation of double stars, but measurement of star diameters with a resolution better than 0.01″ met with unsurmountable difficulties arising when the base D was increased. The optical paths from the two mirrors to the photographic plate must be maintained equal with an accuracy of the order of the wavelength; this condition imposes extraordinary requirements on the accuracy and rigidity of the entire structure. Even greater difficulties are connected with the fluctuations of the atmosphere, especially when the base is long. The difficulties are best illustrated by the fact that the interferometer constructed at Mount Wilson, with a base of 15 meters, yielded no reliable results.

At the same time, greatest interest attaches to the measurement of the angular stellar diameters appreciably smaller than 0.01″, particular hot stars of the earlier O and B classes and Wolf-Rayet stars. Such measurements are important because knowledge of the angular dimensions yields the effective temperature of the star surface. Estimates have shown that a base 400 meters long is required for a complete resolution of the brightest class O star in the sky; complete resolution of Wolf-Rayet stars is possible only with bases longer than 2 kilometers. Obviously, there is no hope of attaining such a resolution in Michelson's stellar interferometer.

Michelson's interferometer can be employed with greater success in radio astronomy, because the resolution of radio telescopes is much smaller than the resolution of optical telescopes. Therefore the use of interferometers in radio astronomy offers a much larger relative gain in resolution, although the maximum resolution of modern radio interferometers is inferior to the resolution of Michelson's stellar interferometer. The maximum resolving power of a modern radio interferometer is 1 second of arc (50 times worse than that of the stellar interferometer), and the corresponding resolution of a radio telescope is 2,000 times worse than the resolution of an average optical telescope. In the middle 50's, the radio interferometer was the principal instrument of radio astronomy.

The construction of a radio interferometer differs somewhat from that of a stellar interferometer, although the operating principle is the same. The interferometer consists of two antennas with amplifiers, spaced a distance D apart (Fig. 2). Signals from two points are guided to the observation point O, where the interference pattern is observed either by varying the phase of one of the signals or as the result of the motion of the radiation source relative to the base (due to the earth's rotation about its axis). The signals are fed to point O through either a high-frequency cable or a radio-relay line. In the latter case it is necessary to use frequency conversion, since it is impossible to obtain a high gain at a single frequency (a high gain is essential to eliminate additional noise pickup along the radio-relay path). The frequency conversion involves in turn the necessity of transmitting along the radio relay line not only the converted signal, but also the phase of the heterodyne used for the conversion.

As already indicated, the greatest resolution attained so far in a radio interferometer is 1″. This is the radio interferometer of the Jodrell-Bank observatory (England), with a base of approximately 132 kilometers, operating at a wavelength of 73 cm [6]. Measurements with this radio interferometer have shown that several per cent of the several hundred radio sources observed at 73 cm in the northern sky have angular dimensions smaller than 1″. There are grounds for assuming that at shorter wavelengths the percentage of "pointlike" radio sources among the weaker sources is much higher. Figure 3 shows the radio spectrum of a source designated 1934-67 in the catalogs [7] and located in the southern hemisphere. On the basis of the characteristic break in the radio spectrum at a frequency close to 1,000 Mcs we can expect its angular dimension to be of the order of 0.005″. Of greatest interest in the investigation of "pointlike" radio sources is the fact that it is precisely among them that objects of unusual nature, such as quasars, are observed. It is not excluded that the observation and a detailed study of "pointlike" radio

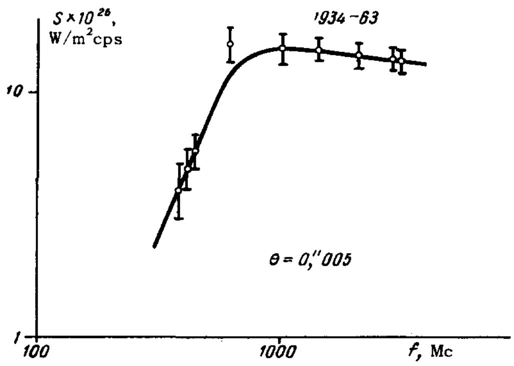

FIG. 3. Radio spectrum of 1934-63 source with small angular dimension.

sources will lead to a discovery of new and hitherto unknown objects in the universe. In particular, hypothetical transmitters from extraterrestrial civilizations should be received on earth as "pointlike" radio sources [8].

It is difficult to increase the resolution of radio interferometers in their modern form to hundredths or thousandths of a second, for the same reasons as in the case of stellar interferometers. The main obstacles to an increase in the base length is the instability of the phase along the communication lines between the points and the observation point O. Fortunately, in a radio interferometer it is possible to eliminate these instabilities, so that interferometers with a base length limited only by the diameter of the earth's sphere (that is $D \sim 13{,}000$ km) are feasible. (This corresponds to a resolution of approximately $0.001''$ at 10 cm wavelength.) They are made possible by eliminating the transmission of the signal and of the heterodyne phase over communication lines [9], separately recording the signals on a magnetic tape at the two points, shipping the tapes to a single point, and then processing the signals simultaneously as with an ordinary radio interferometer. As to the heterodyne phases, they can be made sufficiently stable with the aid of modern atomic and molecular generators, so that the phase difference remains constant during the time of observation. Calculations show that for a radio interferometer with a transcontinental base the required heterodyne stability is of the order of 10^{-10}. The construction of such an interferometer is shown in Fig. 4. The convenience of such a scheme is obvious. It is possible in principle to construct a global system of radio interferometers, incorporating all the major radio telescopes of the world, by equipping them with atomic generators and with magnetic-recording equipment. The resultant records can be processed anywhere on earth. The principal difficulty is that all the independent reading must be referred to a single time. In fact, when signals are recorded in a band B it is necessary that the relative time shift between the individual records not exceed the correlation time, which is approximately equal to $1/B$. If $B = 1$ Mc, then the relative time shift must not exceed 10^{-6} sec, whereas modern time services are capable of comparing two remote clocks accurate to about 10^{-3} sec, a figure primarily governed by the uncertainty in the propagation time of the radio signals with which this comparison is made. This creates a paradoxical situation wherein we have at our disposal atomic or molecular clocks accurate to better than 10^{-10}, but we cannot check one against the other with accuracy better than 10^{-3} sec, owing to the instability in the radio signal propagation velocity! Thus, while striving to eliminate these instabilities by foregoing the communication line, we again are faced with it in a somewhat different aspect. There are several methods of overcoming this difficulty. The most obvious is to transport the clocks to a single place where they can be verified, and then return the clocks to the observatory. In order for such comparisons not to be too frequent (say once a year), the stability of the clocks must be increased by five or six orders of magnitude, which is apparently beyond the capabilities of modern technology. It is interesting that such a stability calls for inclusion of relativistic corrections for the airplane speed (if the clock is shipped by air) and for the difference in the gravitational potential at different points on earth. It is also possible to check the clocks with the aid of special artificial satellites. In principle we can get along with the existing time service. This calls for breaking up the 10^{-3} sec uncertainty interval into 1000 time intervals of duration equal to the correlation time (10^{-6} sec)

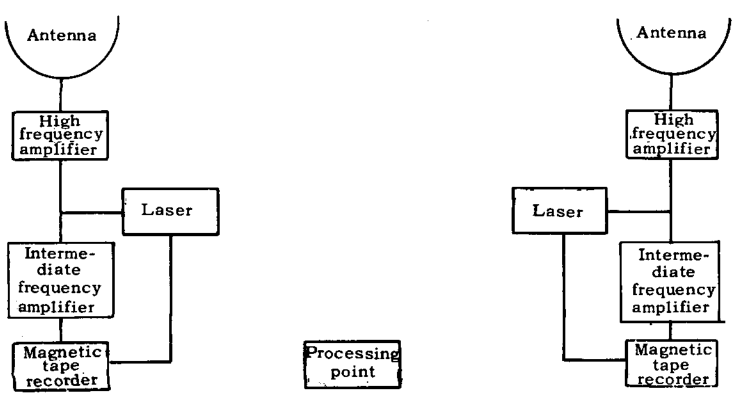

FIG. 4. Radio interferometer with magnetic tape recording.

and obtain one thousand different records successively shifted by 10^{-6} sec. We then chose the particular record for which correlation is observed. This work, of course, must be automatized with electronic computers.

3. INTENSITY INTERFEROMETER

We have seen that the stellar interferometer of the Michelson system has practically reached the "ceiling," that is, the limit of its abilities with respect to resolution, back in the 30's. On the other hand, radio interferometers based on the same principle are being continuously improved and their resolution will probably exceed that of Michelson's stellar interferometer within the nearest few years. Nonetheless, it is just in radio astronomy that a new interferometer principle was first advanced, leading to an appreciable increase in resolution of the optical interferometer. In 1949, before any of the radio sources were resolved (even the most powerful A-Cassiopeiae), the predominant theory of cosmic radiation was the so called radio-star hypothesis. It was assumed that the angular dimensions of the discrete radio-emission sources should be comparable with those of visible stars. The interferometers required for their resolution would therefore have bases of thousandths of kilometers. A Michelson radio interferometer with such a base was not realizable at that time, so that a different interferometry principle, free of the shortcomings of Michelson's interferometer, had to be found. The famous British radio astronomer Hanbury-Brown proposed a new type of interferometer, known as the intensity interferometer [10]. In this interferometer there is no need for coherent heterodynes in two receivers, since it makes use of the coherent property of the noise spectrum obtained after quadratic detection of a high-frequency signal. Such an interferometer is therefore immune to phase fluctuations occurring in the atmosphere or in the ionosphere, and can be constructed with a very short base. In 1951 a model of such an interferometer was successfully tested at 125 Mc in measurements of the distribution of the brightness over the sun's disc. An interferometer having a 3.9 km base (north-south direction) was subsequently used at the same frequency to measure the angular dimension of the two most powerful radio sources, A-Cassiopeiae and A-Cygni, whose diameters were found to be larger than 2′ and of the order of 34″, respectively [11]. A block diagram of the interferometer is shown in Fig. 5 [12]. The antennas used were two arrays of 320 half-wave dipoles, each measuring 40 × 13 meters. To change the base length it was necessary to dismount one 32-element antenna, move it to a new place, and remount it. Receivers with independent heterodynes with 200 kcs bandwidth were used. The video frequencies past the detector were amplified in a frequency band from 1 to 2.5 kcs. The video frequency

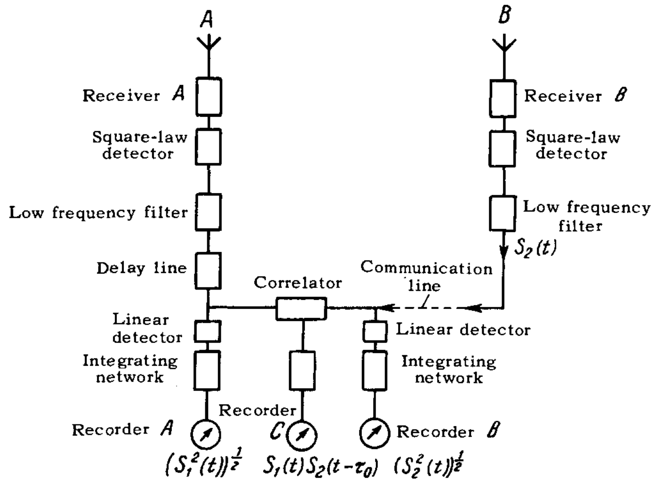

FIG. 5. Intensity radio interferometer.

of one of the receivers was transmitted over a radio line at 83 Mc to the second station, while the frequency of the second receiver was correspondingly delayed by means of a delay line. The outputs of both receivers S_1 and S_2 were then fed to a correlator, in which the product $S_1(t) \times S_2(t)$ was produced, time-averaged in an integrating network, and recorded. Two other recorders fixed the average value of the power $[\overline{S_1^2(t)}]^{1/2}$ and $[\overline{S_2^2(t)}]^{1/2}$. From these three recorded quantities one can calculate the normalized correlation coefficient

$$\varrho^2 = \frac{\overline{S_1(t) \times S_2(t)}}{\{[S_1^2(t)]^{1/2} - P_{N_1}\}\{[S_2(t)]^{1/2} - P_{N_2}\}}, \tag{1}$$

where P_{N_1} and P_{N_2} are the outputs averaged over the time during which the source is outside the antenna directivity pattern.

The normalized correlation coefficient is connected with the distribution of the brightness over the source in the following manner:

$$\varrho^2 = \frac{F_{\cos}^2\left(\frac{2\pi D}{\lambda}\right) + F_{\sin}^2\left(\frac{2\pi D}{\lambda}\right)}{F_{\cos}^2(0)}, \tag{2}$$

where

$$F_{\cos} = \int_{-\pi}^{\pi} P(\theta)\cos\left(\frac{2\pi D\theta}{\lambda}\right) d\theta,$$

$$F_{\sin} = \int_{-\pi}^{\pi} P(\theta)\sin\left(\frac{2\pi D\theta}{\lambda}\right) d\theta$$

are the Fourier components of the angular distribution of the brightness distribution over the source, D is the base length, and λ is the wavelength. Thus, by measuring ρ^2 with different base lengths we can determine both the effective angular diameter of the source and the brightness distribution.

Figure 6 shows the measured correlation coefficient ρ^2 as a function of the base length for the radio source A-Cygni at 125 Mc [13]. The maximum base length in

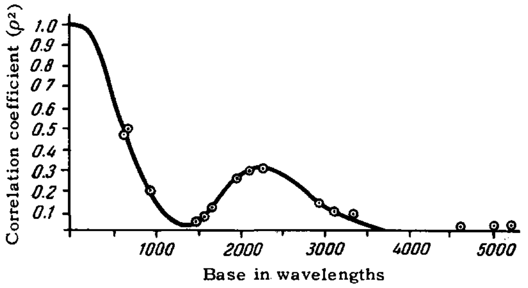

FIG. 6. Dependence of the correlation function on the base length for the radio source A-Cygni.

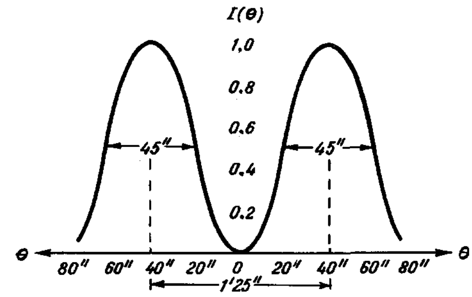

FIG. 7. Model of A-Cygni radio source, according to the results of the measurements with the aid of the intensity interferometer.

these measurements was 5300 wavelengths, corresponding to a resolution of 30″. The curve of Fig. 6. is the calculated dependence of the correlation coefficient ρ^2 for the A-Cygni brightness distribution model shown in Fig. 7. We see from this figure that the source A-Cygni consists of two components. The presence of the secondary maximum at 2,000 λ on the curve of Fig. 6 is an indication of the duality of the radio source.

The same interferometer was used to obtain the brightness distribution over the radio source A-Cassiopeiae.

The use of the intensity interferometer in radio astronomy was limited to these two most powerful radio sources. As expected, the intensity interferometer was insensitive to ionospheric and other phase fluctuations, so that its resolution can be increased almost indefinitely. Subsequently, however, the radio astronomers refocused their attention on shorter wavelengths, in the decimeter and centimeter bands, where ionospheric fluctuations are less pronounced. At decimeter and centimeter wavelengths, most radio sources are much weaker than at meter wavelengths, so that the problem of sensitivity comes to the forefront. In this respect the intensity interferometer is much inferior to the Michelson interferometer. It follows from Eq. (1) that the ratio of the signal at the output of the correlator to the rms uncorrelated-noise fluctuations P_{N_1} and P_{N_2} (which include the intrinsic noise of the receivers and the background noise) is proportional to $(P_S/P_N)^2$, provided P_S (the signal power) is much smaller than P_N (noise power) [10]. For the Michelson interferometer, at the same time, the analogous ratio is proportional to the first power of P_S/P_N. Thus, the signal/noise ratio of the intensity interferometer is worse when the measured signal is weaker than the noise. For example, if $P_S/P_N = 0.01$ (as is quite common in astronomy), at the output of the intensity interferometer this ratio drops to 0.0001, i.e., measurement becomes impossible. If we assume that reliable measurements can be carried out at an output signal/noise ratio of 1/100, then the use of the intensity interferometer is advantageous only for sources which give a signal/noise ratio $\geq 1/10$ at the antenna, thus limiting its use, at the present state of radio astronomy, to only several dozen sources whose angular dimensions are already known.

From the point of view of the technical realization, the intensity interferometer is much simpler and cheaper than the Michelson interferometer. In particular, the requirements with respect to the radio relay communication lines and the magnetic recording systems, described above for the Michelson interferometer, become much simpler, for in this case there is no need for storing information on the phase of the high-frequency signal. The communication between observatories separated by large distances can be effected in principle over existing television relay lines, such as Eurovision and Intervision, or via communication satellites such as Telstar, whereas for the Michelson interferometer it is necessary to construct special radio relay lines in which the phase of the high frequency signal is maintained constant. In the future, therefore, once radio telescopes that provide a larger signal to noise ratio even for the weakest of the presently known sources become available, the investigation of the angular dimensions of these sources (if still unresolved by the future radio telescopes) will be best carried out with the aid of an intensity interferometer.

An approximately similar situation has developed by now in optical astronomy, where even the brightest stars cannot be resolved by the largest telescopes, and the possibilities of the Michelson interferometer have turned out to be too limited. The same Hanbury-Brown, who developed the operating principle of the intensity interferometer in radio astronomy, advanced in [10] the hypothesis that the angular dimensions of stars can be measured in the optical band with the aid of an interferometer based on the same principle as the intensity radio interferometer. The construction of the optical interferometer in this case simply duplicates the construction of the radio interferometer (Fig. 8). The role of the antennas is assumed by parabolic mirrors, and the quadratic detector is replaced by photomultipliers in which, as is well known, the current is proportional to the photocathode illumination. Filters and amplifiers separate the low-frequency

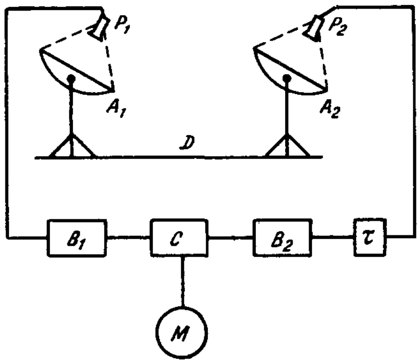

FIG. 8. Stellar intensity interferometer.

fluctuations of the photomultiplier current, which are then fed to a correlator similar to that of the radio interferometer.

The existence of correlations between the fluctuations of the illumination of two photomultipliers on which light from a single source is incident was observed in an experiment especially set up by Hanbury-Brown and Twiss [14]. They determined the ratio of the readings at the output of a correlator S_l to the rms fluctuations at the output of a correlator N_l for two cases: optically aligned cathodes, and cathodes separated by 1.8 cm. It turned out that in the case of aligned cathodes this ratio was +4.2—+7.6 (theoretically S_l/N_l = +5.2—+8.4), whereas for separated cathodes S_l/N_l = −0.4—+1.7 as against a theoretical zero value. This demonstrated the good agreement between the experimental results and the theoretically predicted existence of correlation between the photons in two coherent (that is, coming from a common source) rays of light.

This report evoked a rather lively discussion. The experiment was repeated by others with somewhat different procedures. The correlator was replaced by a coincidence counter for pulses from two photomultipliers, and the number of obtained coincidences was compared with the expected number of random coincidences [15,16]. These experiments disclosed no deviation of the number of coincidences from the expected number of random coincidences, which led the authors of [15] to the conviction that there can be no correlation between the instants of incidence of the photons on the photocathode, determined from the instant of knocking-out of the photoelectron. It was even stated that the existence of such a correlation would call for a radical review of some fundamental principles of quantum mechanics. Hanbury-Brown, Twiss [17,18], and Purcell [19] have shown that the results of these *experiments under the conditions, under which they* were carried out, do not contradict the theory. In response to a criticism by Fellgett, [21] they have set up a new experiment using a coincidence counter under suitable conditions and demonstrated the good agreement [20] with their theory and with the principles of quantum mechanics and thermodynamics.

4. SPATIAL CORRELATION OF RADIATION

The occurrence of discussions on a question that has no bearing on any new and still unknown physical phenomena is due to the relative complexity and even subtlety of the operating principle of the intensity interferometer. At first glance, the possibility of a correlation between the envelopes of signals received in spatially-separated points is highly doubtful. An even greater misunderstanding is due to the statement that the resolution of the intensity interferometer is equal to λ/D, as in the Michelson interferometer. In fact, any random stationary process $\xi(t)$ (and most types of natural radiation can be assumed with great degree of accuracy to be stationary random processes) can be represented in the form [22]

$$\xi(t) = E(t)\cos\Phi(t), \tag{3}$$

where

$$E(t) = \sqrt{\xi^2(t) + \eta^2(t)},$$
$$\Phi(t) = \tan^{-1}\frac{\eta(t)}{\xi(t)},$$

and

$$\eta(t) = -\frac{1}{\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\xi(t)}{\tau - t}\,d\tau$$

is a stationary random process conjugate to $\xi(t)$ and obtained from $\xi(t)$ by Hilbert transformation. The random processes $E(t)$ and $\Phi(t)$ defined in this manner are called respectively the envelope and the phase of the random process $\xi(t)$. If the random process $\xi(t)$ is narrow-band, that is, if its energy spectrum is concentrated in a relatively narrow band about a central frequency ω_0 (these are precisely the processes which obtained in most measuring instruments), then

$$\xi(t) = E(t)\cos[\omega_0 t - \varphi(t)]. \tag{4}$$

In this case the envelope $E(t)$ and the phase $\varphi(t)$ are random functions that vary slowly in time compared with $\omega_0 t$. It can be shown that the characteristic time of variation of $E(t)$ and $\varphi(t)$ is of order of $1/B$, where B is the band width.

In a linear or square-law detector, only the envelope or its square are separated. In particular, in the intensity radio interferometer (see Fig. 5) the output of the square-law detector is of the form (4). The function of the square-law detector is to square $\xi(t)$ and to separate the slowly varying component

$$[\xi(t)]^2 = \frac{1}{2}E^2(t)\{1 + \cos 2[\omega_0 t - \varphi(t)]\}.$$

The second term is rejected by a filter, leaving only

$$[\xi(t)]^2_{\mathrm{LF}} = \frac{1}{2}E^2(t)$$

which is the slowly varying component.

Therefore at the output of a square-law detector there is no information whatever on the phase and on the frequency ω_0 of the input signal. How then is the resulting power determined by the ratio of the wavelength λ to the base D? After all, the correlator receives only signals $[\xi(t)]^2_{LF}$, which do not contain the frequencies ω_0 ($\omega_0 = 2\pi c/\lambda$)! We shall show that if the radiation source has a finite dimension, information concerning the frequency ω_0 is contained also in the envelope. To this end we must consider the manner in which the envelope is formed in the case of a source of finite dimension. Let the source be perpendicular to the base: its brightness is uniform within limits $\pm\theta/2$ and is equal to zero when $|\psi| > \theta/2$ ($\theta \ll 1$). A source element $\Delta\psi_k$ produces at one point of the base a component

$$\xi_k\left(t-\frac{D\psi_k}{2c}\right)=a_k\left(t-\frac{D\psi_k}{2c}\right)\times\cos\left[\omega_0\left(t-\frac{D\psi_k}{2c}\right)-\varphi_k\left(t-\frac{D\psi_k}{2c}\right)\right],$$

and at the other point

$$\xi_k\left(t+\frac{D\psi_k}{2c}\right)=a_k\left(t+\frac{D\psi_k}{2c}\right)\times\cos\left[\omega_0\left(t+\frac{D\psi_k}{2c}\right)-\varphi_k\left(t+\frac{D\psi_k}{2c}\right)\right].$$

From the entire source we obtain accordingly

$$\sum\xi_k\left(t-\frac{D\psi_k}{2c}\right) \text{ and } \sum\xi_k\left(t+\frac{D\psi_k}{2c}\right), \tag{5}$$

where ψ_k runs through values from $-\theta/2$ to $+\theta/2$. The Michelson interferometer multiplies these two quantities, separates the slowly-varying components, and averages, yielding

$$U=\overline{\sum\xi_k\left(t-\frac{D\psi_k}{2c}\right)\xi_l\left(t+\frac{D\psi_l}{2c}\right)}.$$

Since the radiation from each element of the source is independent of the radiation from other elements, we have $\overline{a_k a_l} = 0$ and

$$U=\frac{1}{2}\overline{\sum a_{k_1}a_{k_2}\cos\left[\omega_0\frac{D\psi_k}{c}-(\varphi_{k_1}-\varphi_{k_2})\right]}.$$

The indices 1 and 2 pertain to the instants of time $t-(D\psi/2c)$ and $t+(D\psi/2c)$.

The envelope a_k and the phase φ_k of a narrow-band random process are slowly varying functions of the time, so that we can put

$$a_k\left(t-\frac{D\psi_k}{2c}\right)a_k\left(t+\frac{D\psi_k}{2c}\right)\approx a_k^2(t),$$
$$\varphi_k\left(t-\frac{D\psi_k}{2c}\right)-\varphi_k\left(t+\frac{D\psi_k}{2c}\right)\approx 0;$$

we then obtain

$$U=\frac{1}{2}\overline{\sum a_k^2(t)\cos\omega_0\frac{D\psi_k}{c}}=\frac{1}{2}\sum\overline{a_k^2(t)}\cos\omega_0\frac{D\psi_k}{c}.$$

$\overline{a_k^2(t)}$ does not depend on k, since the brightness of the source is uniform within the limits $\pm\theta/2$. Taking $\overline{a_k^2(t)}$ outside the summation sign and replacing the sum by an integral, we obtain

$$U=\frac{1}{2}\overline{a^2(t)}\int_{-\theta/2}^{\theta/2}\cos\omega_0\frac{D\psi}{c}\,d\psi\approx\frac{\overline{a^2(t)}\,\theta c}{D\omega_0}, \tag{6}$$

or

$$U=\overline{a^2(t)}\frac{\theta\lambda}{D}. \tag{7}$$

Let us consider now the processing of a signal in an intensity interferometer. In this case the signals (5) are each squared individually, the high frequency components are discarded, and the signals are then multiplied. As the result of squaring we obtain in place of (5)

$$\sum_k\xi_k^2+\sum_{\substack{k,l\\k\neq l}}2\xi_k\xi_l.$$

The first sum yields terms of the form $[1+\cos(2\omega_0 t+\psi)]$, which are of no interest to us, since they contain only the dc component and the double-frequency component. The second term gives a sum of terms in the form

$$a_k a_l\,[\cos\beta+\cos(2\omega_0 t+\delta)],$$
$$\text{where }\beta=\frac{\omega_0 D}{2c}(\psi_k-\psi_l)-(\varphi_k-\varphi_l).$$

This is precisely what we need, since β varies slowly in time and contains the term $\omega_0 D/2c$, which determines the resolution of the interferometer. The double-frequency component is again discarded. Taking the foregoing into account, the signal of interest to us at the output of the square-law detector can be written in the form

$$\sum_{k,l}a_k a_l\cos\left[\frac{\omega_0 D}{2c}(\psi_k-\psi_l)-(\varphi_k-\varphi_l)\right]. \tag{8}$$

We see from (8) that the signal at the output of the square-law detector actually contains information on the frequency of the received radiation in the case when $\psi_k \neq \psi_l$, that is, for sources with finite angular dimensions.

We can now find the voltage at the output of the intensity interferometer, caused by multiplying the signals (8) from both points and averaging in the correlator:

$$U=\frac{1}{2}\overline{\sum a_{k_1}a_{k_2}a_{l_1}a_{l_2}\left\{\cos\left[\frac{\omega_0 D}{c}(\psi_k-\psi_l)-(\varphi_{k_1}-\varphi_{k_2})+(\varphi_{l_1}-\varphi_{l_2})\right]+\cos[(\varphi_{k_1}+\varphi_{k_2})-(\varphi_{l_1}+\varphi_{l_2})]\right\}}. \tag{9}$$

The indices 1 and 2 pertain to the instants of time $t-(D\psi/2c)$ and $t+(D\psi/2c)$. Since a and φ are slowly varying functions, we can put $a_1 \approx a_2$ and $\varphi_1 \approx \varphi_2$:

$$U=\frac{1}{2}\overline{\sum a_k^2a_l^2\left\{\cos\left[\frac{\omega_0 D}{c}(\psi_k-\psi_l)\right]+\cos 2(\varphi_k-\varphi_l)\right\}}$$
$$=\frac{1}{2}\sum\overline{a_k^2a_l^2}\cos\left[\frac{\omega_0 D}{c}(\psi_k-\psi_l)\right]+\frac{1}{2}\sum\overline{a_k^2a_l^2\cos 2(\varphi_k-\varphi_l)}.$$

Since the phases φ_k and φ_l are independent, the mean value of the second component vanishes, and $\overline{a_k^2a_l^2} = [\overline{a^2(t)}]^2$.

In analogy with the first case, we obtain ultimately

$$U = [\overline{a^2(t)}]^2 \frac{\theta\lambda}{D}. \tag{10}$$

It follows from this expression that the resulting power of the intensity interferometer coincides with the resulting power of the Michelson interferometer. An important difference is that (10) contains as a factor the square of the mean value of $a^2(t)$, that is, the square of the signal power. This is of fundamental significance for the determination of the sensitivity. In the Michelson interferometer the signal-to-noise power ratios at the input and output are connected by a linear relation. To the contrary, in the intensity interferometer these ratios are connected quadratically, therefore if $S/N \ll 1$, then the signal/noise ratio at the output will be proportional to $(S/N)^2$, in accordance with (10), thus causing the aforementioned great reduction in sensitivity to weak signals.

Another important difference between the two interferometers can also be deduced from the formulas given above. We are referring to the apparatus and atmospheric fluctuations of the signal phase. Because of these fluctuations, φ_1 and φ_2 are no longer equal; their difference varies in time, and in the case of the Michelson interferometer the argument of the cosine in (6) therefore acquires an additional fluctuating phase. If the phase fluctuations exceed 180°, then the mean value of the cosine will be equal to zero and $U = 0$, that is, the interferometer operation becomes impossible. In the analogous expression (9) for the intensity interferometer, the difference $\Delta\varphi_k - \Delta\varphi_l$ is added to the argument of the cosine. It is obvious that this difference is exceedingly small even for very large phase fluctuations. In fact, in the apparatus $\Delta\varphi_k = \Delta\varphi_l$, and in the atmosphere their difference is negligibly small, since the angular dimensions of the source are $\theta \ll 1$. It follows from the foregoing that the intensity interferometer is practically immune to the phase fluctuations that limit the use of the Michelson interferometer.

In radio astronomy the sensitivity of the intensity interferometer is limited by the intrinsic noise of the apparatus and, to a lesser degree, by the background noise. Therefore the use of low-noise amplifiers allows measurement of weaker sources. On going over to shorter wavelengths, for example to the optical band, the situation is radically changed. Most of the noise comes then from the quantum fluctuations of the signal itself, which in principle cannot be eliminated, so that no improvement of the apparatus can increase the sensitivity. Compared with the radio band, the quantum energy in optics is 5 orders of magnitude larger, therefore for an equal energy flux the number of quanta incident on the receiver will be five orders of magnitude smaller; this increases correspondingly the relative dispersion of the number of quanta.

In connection with this peculiarity, it is of interest to consider the operation of an intensity interferometer made up of such purely "corpuscular" instruments as quantum counters and coincidence counters. We have already spoken of an experiment of this kind and of the discussions connected with it. In this experiment two counters connected for coincidence are illuminated by a single light source. The number of coincidences is counted and compared with the expected number of random coincidences due to the finite resolution time of the counters [15-18]. This experiment is essentially similar to the well known Bothe experiment, in which the corpuscular properties of x-rays are determined. In Bothe's experiment the observed number of coincidences did not exceed the expected number of random coincidences, thus demonstrating that the quanta alternately entered either one counter or the other.

Assume that the source emitting the particles with momentum in the interval (p, $p + \Delta p$) has an angular dimension θ; this means that the momenta of the incoming particles lie in a cone of angle width θ. The particles are registered by two counters A and B, spaced a distance D apart (Fig. 9). Counters A and B are interconnected by a coincidence circuit C, which registers the coincidence of the particles in counters A and B whenever the time difference between the arrival of the particles at counters A and B does not exceed the resolving time T of the coincidences circuit. Assume that the same number of particles, n, is incident on counters A and B per cecond. The classical particles strike counters A and B independently, and the probability of random coincidences is the product of the probabilities of incidence of a particle on each counter during the time T, so that the number of coincidences is $n_c^{rand} = (nT)^2$. Thus, the number of coincidences for classical particles is determined only by the resolution time T and depends neither on the distance between the counters nor on the angular dimension θ. This is why these coincidences are called random.

For quantum particles we must take into account

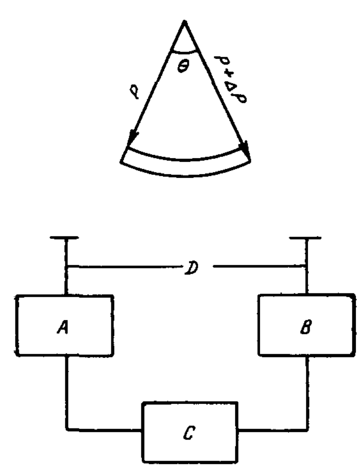

FIG. 9. Intensity interferometer with coincidence counter.

the uncertainty relation and the fact that the particles differ in their statistics from classical particles (Boltzmann statistics). Particles with half-integer spin (electrons, positrons, protons, neutrons, muons) obey Fermi-Dirac statistics, characterized by the property that not more than one particle can be in the same quantum state. Particles with integer spin are described (photons, pions, alpha particles) by Bose-Einstein statistics, whereby the number of particles in each quantum state is not limited. Classical particles have perfectly defined trajectories, so that one particle can always be distinguished from another by its coordinates; classical particles are always distinguishable. In quantum mechanics individual particles correspond to wave packets which occupy a finite volume in space. If the volumes corresponding to the individual particles do not overlap, then the particles are distinguishable and obey classical statistics. The particles whose wave packets overlap are indistinguishable, identical, in the same quantum state, and should therefore obey either Fermi or Bose statistics.

In our experiment with counters, the particles will be indistinguishable if the transverse dimension Δx of the wave packet is larger than the distance between the counters D, while the longitudinal dimension Δz is larger than vT, where v is particle velocity (we confine ourselves to the two-dimensional case for simplicity). The dimensions of the wave packet are given by the uncertainty relations

$$\Delta x = \frac{h}{\Delta p_x}, \quad \Delta z = \frac{h}{\Delta p_z}$$

If $\theta \ll 1$ and $\Delta p \ll p$, then $\Delta p_x \approx p\theta$ and $\Delta p_z \approx \Delta p$.

In our experiment the conditions under which the particles are identical are

$$\frac{h}{\Delta p_x} > D, \quad \frac{h}{\Delta p_z} > vT,$$

or

$$\frac{h}{pD} > \theta, \tag{11}$$

$$\frac{h}{\Delta p \cdot v} > T. \tag{12}$$

for photons $p = h\nu/c$, $v = c$, and conditions (11) and (12) assume the well known form

$$\frac{c}{\nu D} = \frac{\lambda}{D} > \theta, \tag{11'}$$

$$\frac{1}{\Delta\nu} > T, \tag{12'}$$

the first of which is the known formula for the resolution of the interferometer (cf. (7) and (10)), and the second is the condition for the signal to the "narrow-band," a condition essential for interferometer operation.

Particles satisfying conditions (11) and (12) are in the same quantum state (it is assumed that all the particles are polarized; in the case of electrons, for example, this means that all the electrons have identical spin projections) and consequently obey quantum statistics. The remaining particles, which do not satisfy conditions (11) and (12), are distinguishable and can be described by classical statistics. Consequently, out of the n particles incident on each counter per unit time, n_{cl} obey classical statistics and n_{qu} quantum statistics: $n = n_{cl} + n_{qu}$. The relative number of "quantum" particles n_{qu}/n is obviously equal to

$$\delta = \begin{cases} \dfrac{\frac{h}{pD}\,\frac{h}{\Delta pv}}{\theta T} & \text{when } \frac{h}{pD} < 1, \quad \frac{h}{\Delta pv} < 1, \\ 1 & \text{when } \frac{h}{pD} > 1, \quad \frac{h}{\Delta pv} > 1. \end{cases} \tag{13}$$

In the case of Fermi statistics, coincidence of two "quantum" particles is impossible, for not more than one particle can be in one quantum state (Pauli principle). Consequently, the number n_c^F of fermion coincidences (in Fermi-Dirac statistics) will comprise only coincidences of 'classical' particles, the number of which is $n(1 - \delta)$, and coincidences between "classical" and "quantum" particles:

$$n_c^F = [n(1-\delta)T]^2 - \delta nT(1-\delta)nT$$
$$= (nT)^2(1-\delta) = n_c^{rand}(1-\delta). \tag{14}$$

For bosons (Bose-Einstein statistics) we obtain similarly

$$n_c^B = n_c^{rand}(1+\delta). \tag{15}$$

Thus, the number of coincidences of fermions is smaller, and that of bosons larger, than the number of random coincidences of classical particles. This difference between quantum particles and ideal classical gas particles can be described by introducing an effective particle interaction, connected with the quantum-mechanical exchange effects and leading to mutual "repulsion" of fermions or "attraction" of bosons[3]. Owing to the fermion repulsion, the probability of their coincidence in space decreases, leading to a decrease in the number of coincidences in our experiment with counters. The boson attraction, to the contrary, causes them to bunch and increases the probability of coincidences in our experiment. From the foregoing reasoning we can see that quantum effects begin to come into play in those cases when more than one particle can occur in one quantum state, that is, when "degeneracy" sets in. The degree of "degeneracy" can be defined by the parameter

$$\delta = \frac{h^2}{p\Delta p \cdot D\theta T}.$$

If $\delta \ll 1$, then practically all the coincidences will be random. With increasing δ, the number of coincidences will deviate from random in one direction or another, depending on the type of particle. It is natural that this property of the parameter δ can be used to measure one of the physical quantities contained in it (p, Δp, v, D, θ, T, h). In particular, the interferometer measures the angular dimension θ, and the parameter δ is varied by varying the base D.

1. Photons

For photons we have

$$\delta = \frac{\lambda}{D\theta} \frac{1}{\Delta\nu \cdot T}. \tag{16}$$

Consequently, when measuring the angular dimension θ the necessary condition for normal operation of the interferometer is a sufficiently large value of $1/\Delta\nu \cdot T$, that is, a relatively narrow band and a high resolution of the coincidence circuit. With increasing frequency, $\Delta\nu$ increases and therefore the use of the interferometer becomes difficult. In such measurements, the random coincidences play the role of "noise" and the quantum coincidences are the "signal." The random coincidences are determined by classical (i.e., corpuscular) properties of the particles, and the quantum coincidences by their wave properties. Naturally, the higher the frequency, the shorter the wavelength and the smaller the role played by the wave properties. In practice such an interferometer can be used in astrophysics only in the visible and infrared regions of the spectrum. For x-rays and gamma rays, the counters are a convenient means of measurement, but $\Delta\nu$ is so large that the wave properties are lost against the background of the corpuscular ones. In fact, the resolving time of the coincidence circuits is usually not smaller than 10^{-9} sec, whereas the bandwidth in the x-ray region is of the order of 10^{17} sec^{-1}, and consequently $1/\Delta\nu T = 10^{-8}$. In the optical region $1/\Delta\nu T$ can reach unity for very narrow lines. It is possible that exceedingly narrow lines will also be discovered in the x-ray region (for example, of the type of laser lines). Then the interference procedure with coincidence counters can turn out to be very convenient for the measurement of the width of such a line.

2. Electrons

For nonrelativistic electrons p = mv. The angular resolution is determined from (11):

$$\frac{h}{mvD} > \theta \quad \text{or} \quad \lambda/D > \theta,$$

i.e., it is the same as for photons, but the role of the wavelength is played by the deBroglie wavelength h/mv. From (12) we obtain the second condition for the operation of the interferometer

$$\Delta E_k T < h, \tag{17}$$

where ΔE_k is the width of the energy spectrum of the electrons. For the parameter δ we have

$$\delta = \frac{\lambda}{D\theta} \frac{h}{\Delta E_k T}. \tag{18}$$

As before, by determining δ we can measure one of the quantities contained in (18). Analogous relations are valid also for all particles (protons, neutrons, mesons, neutrinos, etc.).

Measurements of this kind are of interest in astrophysics only for uncharged particles, since charged particles moving in the magnetic fields of the galaxy change direction of motion, this being equivalent to an increase in the angle θ. Particles of interest for astrophysics have very high energies and consequently a very short wavelength. This method could therefore be used to measure very small angular dimensions. For example, a neutron star with diameter of the order of 10 km would be seen at a distance of 10 parsec at an angle $\theta = 10^{-8}$ sec. The base required for the resolution of a neutron star in the x-ray region ($\lambda = 1$ Å) is D = 3 km. However, to ensure a sufficient difference between the number of coincidences and the random coincidences it is necessary to receive in a very narrow frequency interval $\Delta\nu$, something not feasible at present because of the weakness of the x-ray sources and the relatively low effective area of the counters. Quasars may have equally small angular dimensions. They may be sources of not only light, x-rays, and gamma rays, but also powerful neutrino emitters. The difficulty in recording neutrinos is well known; it is all the more difficult to record their coincidences. As to neutrons, they have too short a lifetime to traverse galactic or intergalactic distances.

5. MEASUREMENT OF STAR DIAMETERS WITH AN INTENSITY INTERFEROMETER

A detailed exposition of the theory underlying the method of measuring star diameters with the aid of an intensity interferometer, and a discussion of the necessary apparatus parameters, are contained in [24,25]. It is shown in [26] that, owing to the low sensitivity, measurements with an intensity interferometer are confined only to stars visible to the unaided eye. At the same time, the resolution, which is more likely to be determined by limitations of the radio methods than by optical limitations, is sufficient to measure any star with sufficient surface brightness. The operation of the interferometer is not affected by atmospheric fluctuations. Very cold stars with sufficient surface brightness are fully resolved by each mirror of the intensity interferometer, and therefore the method is applicable only to stars of earlier classes than K5. In this case it is possible to employ a modified variant of an interferometer with a single mirror, in which the light flux is split into two parts that are incident on separate photomultipliers. When so modified the interferometer can be used to measure bright stars to class M5.

To check the principle and the calculations, the intensity interferometer was used in the end of 1955 and at the beginning of 1956 to measure the angular diameter of Sirius (α Canis Majoris) [27,28]. The mirrors used were two standard army projectors of 156 cm diameter. A photomultiplier with maximum sensitivity near 4,000 Å was located in the focus of each mirror. The fluctuations of the currents of both photomultipliers were amplified in a bandwidth 5—45 Mc and fed to a correlator consisting of phase modulators, amplifiers, and synchronous detectors. The product of two currents obtained in the correlator and serving as a measure of the correlation between the fluctuations of the light intensities incident on the two photocathodes, was recorded with the revolution counter of an integrating motor M (Fig. 8). A second motor recorded the rms amplitude of the fluctuations at the output of the correlator. This yielded the ratio of the signal to noise at the output.

The signal/noise ratio was measured with a short base (2.56 m) and found to be 8.50 ± 0.67 at a total observation time of 345 minutes. The calculated signal/noise ratio for this base is 9.58, so that the agreement with calculation was satisfactory. At large bases, the signal to noise ratio decreased; amounting to 3.56, 2.65, and 0.83 at 5.35, 6.98, and 8.93 meters, respectively. The measured angular diameter of Sirius, with account of the darkening towards the edge is 0.0071″ ± 0.00055″ and agrees with the theoretical value 0.0069″ ± 0.00044″.

After successfully testing the interferometer model by measuring the angular diameter of Sirius, a larger instrument was designed, intended for the measurement of all stars in the southern hemisphere, of spectral classes earlier than F0, and brighter than photographic magnitude +2.5. The new stellar intensity interferometer has two mirrors of 6.7 m diameter, mounted on an annular rail track 190 meters in diameter, corresponding to a maximum resolution of 0.0005″ [29] The construction of the interferometer is now under completion at the Narrabri (Australia) observatory. In 1963, a series of preliminary observations was made of the bright star Vega (α Lyrae) for the purpose of testing all the elements of the instrument [30]. The results of observations of Vega were the first direct measurements of the diameter of a star of so early a spectral class as A0.

In these observations, the measurements were made in a band 80 Å wide centered at λ = 4385 Å. The fluctuating components of the current of the photomultipliers located at the foci of both mirrors were amplified in a 10—120 Mc band and then multiplied in an electronic correlator. At the output of the correlator, the mean value of the correlation function over 100 seconds was printed in digital form. The observations were carried out with five different bases from 10 to 23 meters (Fig. 10). The total observation time was approximately 28 hours. The plot of the correlation coefficient against the base length, shown in Fig. 10, corresponds, with allowance for darkening towards the edge, to an angular diameter of Vega θ = 0.0037″ ± 0.0002″. The brightness temperature at λ = 4400 Å is equal to 9600 ± 300°K and is in good agreement with the measurement of the color temperature. The radius of Vega is 3.2 times larger than the sun's radius. Comparing the angular diameter of Vega, obtained in

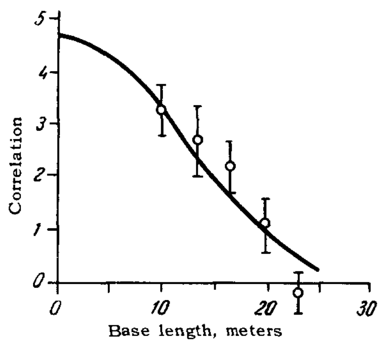

FIG. 10. Dependence of the correlation on the base length for Lyrae.

these measurements, with the smallest angular diameter measured with the aid of the Michelson interferometer we see that the resolving power was increased by almost 10 times. Angular diameters of stars of earlier spectral classes were determined even in the preliminary measurements.

Further efforts towards improving the intensity interferometer can be directed at increasing its sensitivity. For example, to measure the angular diameter of the quasar 3C-273 the sensitivity of the interferometer in Narrabri should be higher by ten stellar magnitudes. An obvious way is to increase the mirror diameter. Another method is to separate bright emission lines and broadening the frequency band of the fluctuations of the photomultiplier current.

The intensity interferometer can be used successfully in radio astronomy to measure relatively powerful radio sources with very small angular dimensions. Another possible application is measurement of angular dimensions of radio sources at long wavelengths from satellites and rockets beyond the limits of the earth's ionosphere.

[1] V. I. Slish, Nature **199**, 682 (1963).

[2] Occultation Method of Determining Star Diameters, UFN **49**, 489 (1953).

[3] C. Hazard, Mont. Not. Roy. Astron, Soc. **124**, 343 (1962).

[4] S. Bowyer, E. T. Byram, T. A. Chubb, and H. Friedman, Science **146**, 912 (1964).

[5] A. A. Michelson and F. G. Pease, Astrophys. J. **53**, 249 (1921).

[6] B. Anderson, W. Donaldson, H. P. Palmer, and B. Rowson, Nature **205**, 375 (1965).

[7] Bolton, Gardner, and Mackly, Nature **199**, 682 (1963).

[8] N. S. Kardashev, Astron zh. **41**, 282 (1964), Soviet Astronomy AJ **8**, 217 (1964).

[9] N. S. Kardashev, L. I. Matveenko, and G. B. Sholomitskiĭ, Izv. Vuzov (Radiofizika) **8**, 641 (1965).

[10] R. Hanbury Brown and R. Q. Twiss, Phil. Mag. **45**, 603 (1954).

[11] R. Hanbury Brown, R. C. Jennison, and M. K. Das Gupta, Nature **170**, 1061 (1952).

[12] R. C. Jennison and M. K. Das Gupta, Phil. Mag. **1**, 55 (1956).
[13] R. C. Jennison and M. K. Das Gupta, Phil. Mag. **1**, 65 (1956).
[14] R. Hanbury Brown and R. Q. Twiss, Nature **177**, 27 (1956).
[15] E. Brannen and H. I. S. Ferguson, Nature **178**, 481 (1956).
[16] A. Adám, L. Jánossy, and P. Varga, Acta Physica Hungarica **4**, 301 (1955).
[17] R. Hanbury Brown and R. Q. Twiss, Nature **178**, 1447 (1956).
[18] R. Hanbury Brown and R. Q. Twiss, Nature **179**, 1128 (1957).
[19] E. M. Purcell, Nature **178**, 1449 (1956).
[20] R. Q. Twiss, A. G. Little, and R. Hanbury Brown, Nature **180**, 324 (1957).
[21] P. B. Fellgett, Nature **179**, 956 (1957).
[22] B. R. Levin, Teoriya sluchaĭnykh protsessov i ee primenenie v radiotekhnike (Theory of Random Processes and Its Use in Radio), Soviet Radio Press, 1957.
[23] L. D. Landau and E. M. Lifshitz, Statisticheskaya fizika (Statistical Physics), Nauka, 1964, page 188.
[24] R. Hanbury Brown and R. Q. Twiss, Proc. Roy. Soc. **A242**, 300 (1957).
[25] R. Hanbury Brown and R. Q. Twiss, Proc. Roy. Soc. **A243**, 291 (1957).
[26] R. Hanbury Brown and R. Q. Twiss, Proc. Roy. Soc. **248**, 199 (1958).
[27] R. Hanbury Brown and R. Q. Twiss, Nature **178**, 1046 (1956).
[28] R. Hanbury Brown and R. Q. Twiss, Proc. Roy. Soc. **A248**, 222 (1958).
[29] H. Messel, Nature **186**, 270 (1960).
[30] R. Hanbury Brown, C. Hazard, J. Davis, and L. R. Allen, Nature **201**, 1111 (1964).

Translated by J. G. Adashko