Measuring the Sizes of Stars

Fringe Benefits of Interferometry

Rajaram Nityananda

Stars, other than the Sun, appear to our unaided eyes as points of light. Large telescopes show an image whose size is dictated by refractive index irregularities in the Earth's atmosphere. The size of this blurring is much greater than that of the star, and hence it is difficult to measure the stellar size. Fizeau showed how one might overcome this limitation using the two-slit interference technique. It was Michelson who carried out this programme and made the first direct measurement of the giant star Betelguse in the constellation of Orion. His value for the angular diameter, 47 milliarcseconds or 2.6×10^{-7} radians, was completely confirmed by later work following his methods. The key concept introduced was 'fringe visibility', which turned out to be very fruitful in the later development of optics as well as astronomy.

Rajaram Nityananda works at the School of Liberal Studies, Azim Premji University, Bengaluru. Earlier, he spent a decade at the National Centre for Radio Astrophysics in Pune, and more than two decades at the Raman Research Institute in Bengaluru. He has taught physics and astronomy at the Indian Institute of Science in Bengaluru and the IISERs in Pune and Mohali. He is keenly interested in optics in a broad sense, and more generally, problems, puzzles, paradoxes, analogies, and their role in teaching and understanding physics.

Keywords
Fringes, interference, spectral lines, interferometer, Fizeau mask, angular size, parallax.

1. Introduction

The nearest star to our solar system, called Proxima Centauri, is about 260,000 times further away from us, than the Sun. The Sun subtends approximately half a degree (roughly 10^{-2} radians) at Earth. Therefore, a star similar to the Sun placed 200,000 times further away would subtend 5×10^{-8} radians. In the more familiar angular units, this is 100 milliarcseconds (one degree = 60 arcminutes = 3600 arcseconds). In popular terms, this is the angle subtended by a cricket pitch of 20 m length on the Earth, as viewed from the Moon. This is just the case of the nearest star. Many of the bright stars we see in the sky could be a hundred times further away than this nearest star. If they were similar to the Sun in size, measuring their angular sizes becomes a very challenging task.

The measurement of the angle subtended by a star on Earth is rather important for astronomy. Knowing the distance, we can multiply the angle subtended in radians by the distance to obtain the physical size of the star.

The measurement of the angle subtended by a star on Earth is rather important for astronomy. Knowing the distance (measured using the parallax method), we can multiply the angle subtended in radians by the distance to obtain the physical size (the diameter) of the star. Data on the spectrum, the mass, and the physical size of a star enable astrophysicists to check their models of stars. This gives confidence in applying the models more widely, even when we cannot directly measure the angular size.

2. The Limitations of Telescopes

The smallest angle θ_{min} which can be measured with confidence by an *ideal* telescope depends on its diameter D and the wavelength of the light used, λ. It is given in radians by $\theta_{\text{min}} = \lambda/D$ (*Figure* 1(a)). For a wavelength $\lambda = 500$ nm in the middle of the visible spectrum, and $D = 1$ m, this formula gives 5×10^{-7} radians or 100 milliarcseconds. It might therefore be barely possible to measure the diameters of nearby stars, somewhat bigger than the Sun, with a one metre telescope.

The largest telescopes available in Michelson's time were not more than about one metre in diameter. He therefore chose the bright red star, Betelguse in the constellation of Orion[1]. This star was known to be about 600 light years away – about 150 times further away than the nearest star we mentioned earlier. It might seem a poor candidate since a distant star would subtend a very small angle. But there was already clear evidence that Betelguse was no ordinary star but a giant. The apparent brightness was greater than nearby stars, despite being much further away, The apparent brightness is proportional to the true brightness, but also to the inverse square of the distance. This means that the true brightness of this star was hundreds of thousands of times greater than the Sun, and this is in spite of being red[2]. All this implies that the star had a very large area – it was a 'giant'. Its diameter might have, in principle, been measured by a large enough telescope.

[1]The name is derived from Arabic for 'the hand of Orion', since this constellation is viewed as a human figure.

[2]A red body emits much less radiation per unit area than a yellow one.

However, there was another fundamental difficulty. Even the best

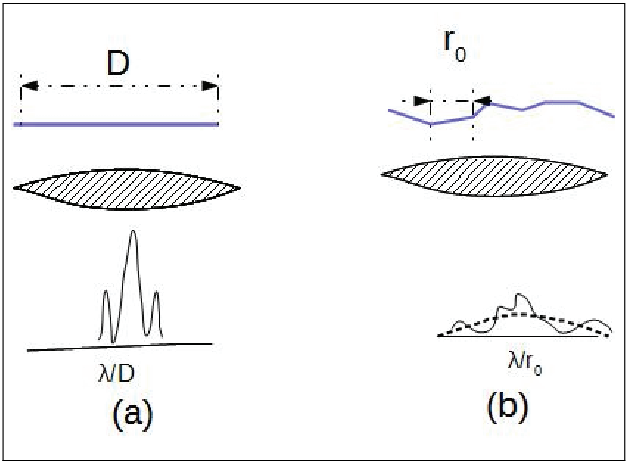

Figure 1. Image of a point source **(a)** Ideal telescope: The diameter of the lens (or mirror) is D, and we have a plane wavefront, shown in blue, falling on it from the point source. The image is a diffraction pattern, with a maximum at the centre and surrounded by weak rings. The image size depends on the focal length, but corresponds to an angle λ/D on the sky **(b)** The wavefront is now randomly distorted by the Earth's atmosphere, as shown by the blue line. Over a size of r_0 (typically 10 cm at visible wavelengths), the error in phase is of the order of a radian. The image consists of a region of (angular) size λ/r_0, with random interference maxima and minima. The atmosphere changes in a short time of the order of 10 milliseconds. So in a few seconds this averages to a smooth image of size λ/r_0 which is about 1 arcsecond. This is shown as a dashed line.

made telescopes do not produce the ideal image of angular size $\theta_{\min}$ which we have discussed above. The reason is explained with more quantitative detail in *Figure* 1(b). To produce a sharp image, light from the distant star must converge at the focus of the telescope. At the peak of the image, light arriving from all parts of the mirror has to be in phase, to produce the maximum intensity. Any errors in the mirror or refractive index variations in the atmosphere, spoil this good phase relationship. As a result, the intensity at the focus goes down. However, since the total energy is the same, namely what fell on the mirror, it simply gets spread over a larger region of the focal plane. In other words, we have a blurred image. Its size for a typical telescope is around one arcsecond or 5×10^{-6} radians. So it is clear that the actual size of the star would be vary hard to find from such a blurred image.

3. Using the Interference of Light

Fizeau, the great French pioneer of optics, came up with a method to partially overcome this difficulty. He proposed and implemented what is now called a 'Fizeau mask'. The idea is to cover the whole telescope with an opaque object, leaving just two holes less than 10 cm in size. Over 10 cm, the path differences caused by refractive index variations in the atmosphere are significantly less than one wavelength (under good conditions!). Therefore,

Box 1. Fizeau's Proposal for Measuring the Size of a Star

There exists indeed for the majority of the phenomena of interference, such as the fringes of Young, those of the mirrors of Fresnel, and those which give place to the scintillation of stars according to Arago, a relation remarkable and necessary between the dimension of the fringes and that of the source of light, so that finely spaced fringes cannot occur when the source of light has anything more than almost insensible angular dimensions; from where it is perhaps allowed to hope that based on this principle, while forming the interference fringes at the centre of the large instruments intended to observe stars, for example, by means of two broad very-isolated slits, it will become possible to obtain some new data on the angular diameters of these stars.

when the light from these two apertures is combined at the focus of the telescope, we get a nice two-slit interference pattern (*Figure* 2). It is true that the path difference between the two apertures could be much more than the wavelength, but that only shifts the centre of the pattern. This path difference also changes with time, because the atmospheric refractive index irregularities are not fixed. They come from temperature variations and are blown by the wind. This means that the entire fringe pattern will move, and the observer has to be quick enough to catch the fringes. But a moving fringe pattern should be regarded as a lesser evil than a blurred image as we will see below. *Box* 1 has a rough translation from the French of Fizeau's original proposal, which is compressed into a single sentence!

He is just saying that closely spaced fringes get washed out if the source of light illuminating the two slits is too large – which was known. The new idea is that this disappearance of interference fringes can be exploited to measure the angular size of the source, even if it is a star, by placing the slits quite far apart in front of a large telescope. This is illustrated and explained qualitatively in *Figure* 2(a) and its caption. The fundamental concept to keep in mind for the rest of this article is as follows.

Real light sources, such as stars, are a combination of multiple monochromatic point sources. These sources are considered 'incoherent' with respect to each other as they do not have any stable phase relationship.

A monochromatic point source – MPS for short – produces interference fringes, when the light from it reaches the detector by different paths. This is an ideal situation. Any real source, such

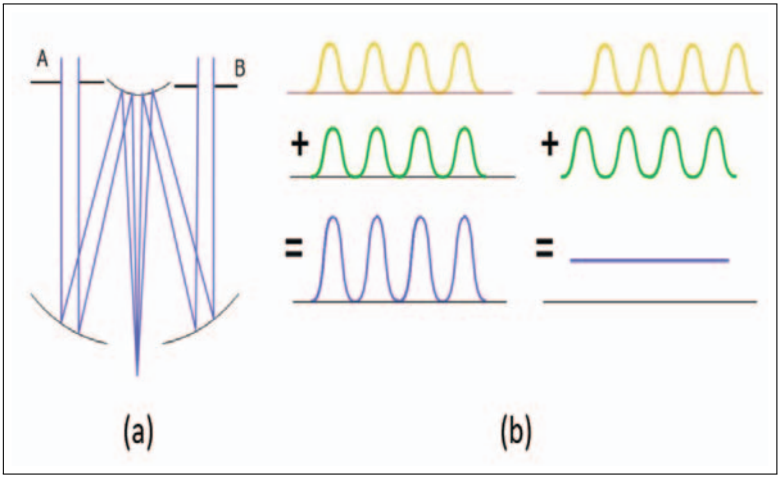

Figure 2. **(a)** The Fizeau mask. Light from a distant point source falls on the telescope's primary mirror, but only through two smaller holes of size 10 cm or less – A and B in the figure. This ensures that the two interfereing wavefronts are approximately plane, in spite of atmospheric disturbances. The blue lines are rays which show the light paths which focus after two reflections at the bottom of the figure. Each of the images would correspond to approximately 1 arcsecond on the sky, and they would overlap to give interference fringes. In practice, the holes need not be at the top but smaller holes nearer the focus would achieve the same effect. **(b)** Fringes at the focus for a binary star with two equal components. Depending on the baseline, we get fringe visibility varying from 1 to 0.

as a star, or a gas of atoms emitting a spectrum, is regarded as a combination of multiple MPS'. These sources, at different locations, and at different frequencies, do not have any stable phase relationship. We call such sources 'incoherent' with respect to each other. We calculate the interference pattern due to a each MPS separately, and add the intensities of the different patterns.

This model of light served for what are called the 'thermal' sources, and only had to be improved when laser sources with long-lasting phase stability, were introduced – but that is another story.

The basic conclusion from *Figure* 2(b) and the discussion in the caption is that when we form interference fringes from a star with two apertures as Fizeau suggested, we get clearly visible fringes when the spacing between the apertures, denoted by b is very small. As b increases, the fringes become weaker. As the separation approaches a value of the order of $b_{\rm max} \simeq \lambda/\Theta$, the fringes disappear. The reason is simply that the fringe patterns due to the different parts of the source have maxima at different locations on the screen. Once these locations are spread out by greater than the fringe width λ/Θ, maxima of some parts of the source overlap with the minima of others. The result is to reduce the variations in intensity. Notice that this mathematical argument is the same as used in discussing the destructive interference between two coherent sources[3]. But do keep in mind that the physics is different

[3]Produced from the same source illuminating two slits.

– we are adding intensities which are positive quantities precisely because the different parts of the source are incoherent.

This general principle is sufficient to appreciate the famous experiment of Michelson and Pease in 1921, to which we now turn. For a more quantitative formulation see *Box* 2.

Box 2. Fringe Visibility for Double and Single Stars

Let us first consider a system of two stars, separated by an angle $\Delta\alpha$, each treated as a monochromatic point source for simplicity. We can think of each star as producing a set of fringes. Because the two stars are not coherent sources, we add the intensities in the two fringe patterns.

If I is the average intensity in a fringe pattern, with spacing x on the screen, the intensity at the point x is described by the function $I(x) = I(1 + \cos(2\pi x/s))$. Notice that this goes all the way from 0 to $2I$.

It is clear that the contrast of the combined fringe pattern will be the best when the maxima of the fringe pattern of one star fall on the maxima of the other. Take the intensities of the two stars to be I_1 and I_2 with $I_1 \geq I_2$. In this case, the combined intensity of the two sets of fringes at a point x on the detector is proportional to $I_1(1 + \cos(2\pi x/s)) + I_2(1 + \cos(2\pi x/s))$, where s is the fringe spacing. Note that we have aligned the maxima of the two cosine α functions. The resulting pattern has a maximum proportional to $2(I_1 + I_2)$, and a minimum of zero. If, however, the maximum of one fringe pattern falls on the minimum of the other, then the total intensity on the detector is given by $I_1(1 + \cos(2\pi x/s)) + I_2(1 - \cos(2\pi x/s))$. Note the minus sign in the second term, which tells us that the second star fringe pattern has a dark fringe at $x = 0$. Now the maximum intensity is given by $2I_1$ (when the cosine is 1) and the minimum by $2I_2$ (when the cosine is -1). Michelson defined fringe visibility as $V = (I_{\max} - I_{\min})/(I_{\max} + I_{\min})$.

This is a rather natural definition. An electrical engineer would call it the ratio of the AC component of the intensity variation, which is (maximum − minimum)/2, to the DC component (maximum + minimum)/2. We are dividing the amplitude of the sinusoidal intensity variations in the pattern by the average intensity. Clearly the maximum possible value is 1. In our example, when the two sets of fringes are aligned, maximum on maximum, we have $I_{\min} = 0$, and hence $V = 1$. When the fringes are displaced by half the fringe width – maximum falling on minimum, we can use the values of $I_{\max} = 2I_1$ and $I_{\min} = 2I_2$ calculated earlier. This gives $V = (I_1 - I_2)/(I_1 + I_2)$. When the two stars are equally bright, the fringe visibility is zero in this case. We can understand this without calculation – two sets of fringes of equal strength but with maximum of one falling on the minimum of the other will add up to a uniform intensity as a function of x – the fringes are invisible.

The condition for the alignment of the two fringe patterns is explained in *Figure* 2 and can be understood as follows. The separation of the centres of the two masks is denoted by b which stands for 'baseline', a term we will use later.

Contd.

Box 2. ***Contd*****.**

A distant source of light on the perpendicular bisector of the two masks, will have zero path difference. Moving the source by a small angle α (much less than one radian) away from this direction will create a path difference $p = b \times \alpha$. So if we have two stars separated by $\Delta\alpha$ the two path differences at the two apertures of the mask will themselves differ by $\Delta p = b\Delta\alpha$. If this is an integer number of wavelengths, then the maxima of the two fringe patterns will fall on top of each other. Note that this is not sensitive to further path differences after the two masks, which are common to both sets of fringes. The actual separation of the fringes on the detector depends on the angle made between the two beams at the detector, which can be chosen by the experimenter, but this again does not affect the condition for alignment of the maxima of the two fringe patterns. So the fringe visibility is really a property of the light wave from the two stars at two points separated by the baseline b.

The more general case is when the two sets of fringes are neither aligned maximum to maximum nor maximum to minimum. We can choose the origin of x to be the maximum of the fringe pattern of one of the stars which is then $I_1(x) = I_1(1 + \cos(2\pi x/s))$. The fringe pattern of the second star is displaced by an amount ϕ, and is given by $I_2(x) = I_2(1 + \cos(2\pi x/s + \phi))$. The displacement ϕ between the two fringe patterns depends on the angular separation of the two stars. The two previous cases which we discussed were $\phi = 2n\pi$ (path differences differ by wavelength times an integer) and $\phi = (2n + 1)\pi$, corresponding to Δp, an odd number of half wavelengths. The general expression is $\phi = 2\pi\Delta p/\lambda = 2\pi b\Delta\alpha/\lambda$.

It is a mathematical exercise to find the maximum and minimum of $I_1(x) + I_2(x)$ and hence calculate the fringe visibility. This turns out to be very simple in the case of a symmetrical double star, $I_1 = I_2$. In this case we can rewrite $I(x) = 2I(1 + \cos(2\pi x/s + \phi/2)\cos(\phi/2))$ using the standard trigonometric identity for $\cos C + \cos D$. Dividing the amplitude of the variation, $2I\cos\phi/2$, by the average $2I$, the visibility at baseline b is just $V(b) = \cos(\phi/2) = \cos(\pi b\Delta\alpha/\lambda)$. This agrees with our earlier result for the two cases $\phi = 2n\pi$ and $\phi = (2n + 1)\pi$. (Note that a visibility of -1 simply means that the fringes have reversed – the maxima have been replaced by minima and *vice versa*)

We now come to the case of a single star, which is no longer treated as a point source, but as a circular disc. From *Figure* A, we see that the star can be split into a number of fictitious double stars whose separation $2r$, goes all the way from 0 to $2R$, where R is the radius of the star. Further, the figure shows that these double stars are symmetric, and the intensity falls off as we increase the separation from 0 to $2R$. The intensity is just given by the area of the strip which is $2\sqrt{(R^2 - r^2)}dr$. Since our expression for the visibility of a double star is in terms of angles, we should set $\Delta\alpha = 2r/D$ where D is the distance to the star.

We can now use our earlier result for the double star to build up the visibility of the fringes produced by a uniformly illuminated circular disc. *Contd.*

---

* **Note:** We have assumed that the intensity is uniform over the disc, so that the amount of light received is proportional to the area of the strip. Further, by making the strip perpendicular to the baseline, we ensure that the phase differences at the two apertures are the same for all parts of the strip – it is only movement along the baseline which generates phase change.

Box 2. *Contd*.

We divide the amplitude of the fringe pattern by the total intensity. The numerator is a sum of the binary star visibility expressions for pairs of strips going from $r = 0$ to $r = R$. The denominator is just the total intensity. The resulting expression for the fringe visibility measured at a baseline b is:

$$V(b) = \frac{\int_0^R I\sqrt{(R^2 - r^2)}\cos(\pi\, br/D\lambda)\mathrm{d}r}{\int_0^R I\sqrt{(R^2 - r^2)}\mathrm{d}r}$$

Changing variables to r/D, it is clear that this is a function of $R/D = \Theta$, the angular diameter of the star. The resultant visibility function is plotted in *Figure* B as a function of b. Notice that the fringe visibility decreases as the baseline b increases, and becomes zero for $b = 1.22\lambda/\Theta$, where Θ is the angular diameter of the star. Unlike in the case of the binary star, the visibility does not rise back to 1, but undergoes damped oscillations. This same function occurs as the diffraction pattern of a circular aperture. As earlier, the negative values mean that the fringes reverse after becoming invisible, maxima appearing where minima were earlier.

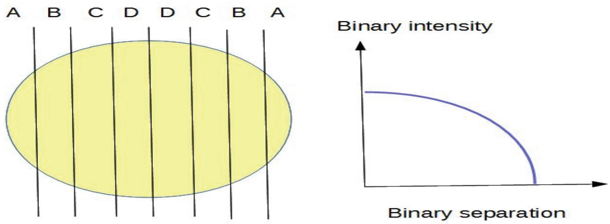

Figure A. Left Disssection of a circular disc into binary pairs of increasing separation and decreasing intensity (D is maximum, A is minimum). **Right** Intensity as a function of separation, given by $\sqrt{1 - (r/R)^2}$.

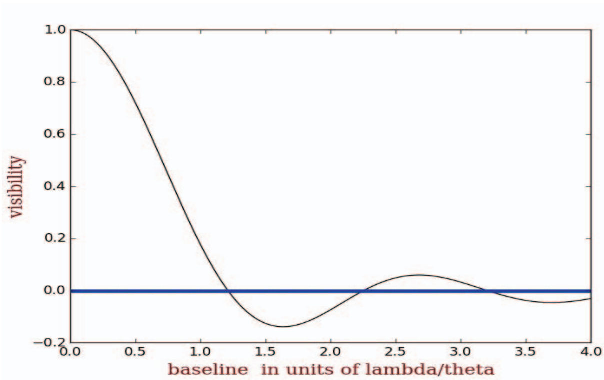

Figure B. Visibility as a function of the observing baseline (mirror separation) for a star, modeled as a uniformly illuminated disc of angular diamter Θ. This plot is based on the equation given in *Box* 2, and shows that the fringes disappear at a baseline $b = 1.22 \times \lambda/\Theta$.

4. The Michelson–Pease Experiment

Michelson used two apertures on a 1.5 m telescope to find the diameters of the moons of Jupiter, as early as 1890. These are somewhat less than 1 arcsecond in angular size (the distance of Jupiter from Earth varies, of course), and hence the baseline needed to achieve zero fringe visibility is about 10 cm. His ambition was to measure the diameters of stars – angles which were much smaller, (by a factor of twenty and more). This clearly needed a larger telescope, and one was completed in 1917 – the 'Hooker telescope' with a 2.5 m diameter primary mirror, at Mount Wilson in the western coast of the United States. The director of the observatory, G E Hale, was persuaded to allow modifications for this experiment to be carried out. Its not clear whether such a permission would be granted by the director of an observatory today; he would worry too much about potential damage!

Michelson used two apertures on a 1.5 m telescope to find the diameters of the moons of Jupiter, as early as 1890. His ambition was to measure the diameters of stars – angles which were much smaller.

Michelson came up with a more ambitious scheme. A large and rigid steel beam, 6 m in length, was made on which plane mirrors could be mounted, and this was fixed on top of the telescope (*Figure* 3). The schematic ray paths show how this simulated the effect of two apertures spaced by 6 m, more than the telescope diameter! The telescope itself provided the mechanism to bring the light beams together where the interference fringes would be observed with an eyepiece.

We quote the crucial paragraph from their paper below. Amateur astronomers will recognise that α Orionis stands for the brightest star in the constellation of Orion, none other than Betelguse. Other stars are named similarly – β Persei is the second brightest star in Perseus.

The 'zero fringe' refers to a separate image with a small baseline where one expects to see fringes on all stars if the instrument and atmosphere are behaving properly.

> *On Dec 13, 1920, after preliminary settings on β Persei with the mirrors separated by 81 inches (229 cm) and on β Persei, and γ Orionis with a separa-*

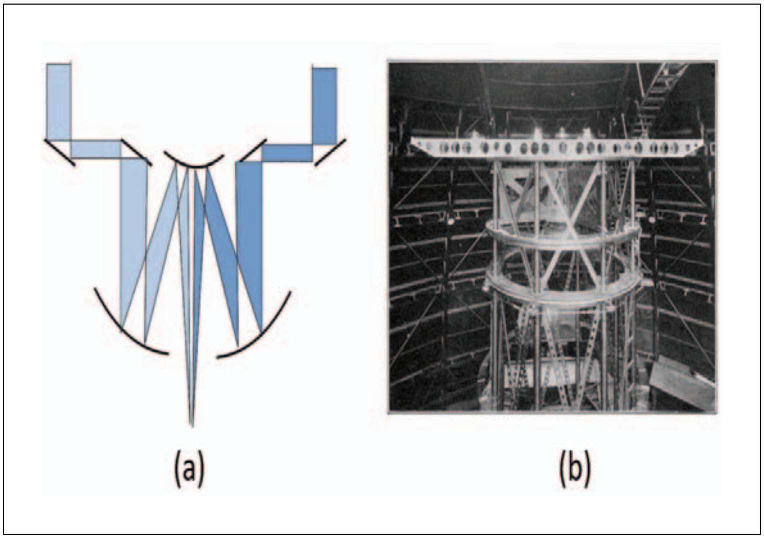

Figure 3. **(a)** Schematic ray paths for the Michelson stellar interferometer. **(b)** The Hooker – also known as the 100 inch telescope at Mount Wilson observatory, with the beam carrying the mirrors mounted on top. Image courtesy: George Ellery Hale – The New Heavens, Public Domain, https://commons.wikimedia.org/w/index.php?curid=1223713

> *tion of 121 inches, thus insuring that the instrument was in perfect adjustment, it was turned on α Orionis and fringes across the interferometer image were sought for some time, but they could not be found. The seeing was very good and the zero fringes could be picked up at will. When next turned on α Canis Minoris, the fringes stood out on both images with practically no adjustment of the compensating wedge,* ***which furnishes a check on the disappearance of the fringes for α Orionis****.*

We have placed the last sentence in bold – the final measurement of the angular diameter depends on this. The beauty is that this did not require quantitative measurement of the fringe visibility, which would have been very difficult at that time. This reminds us of the great detective Sherlock Holmes using the incident of the dog that did not bark in the night time to solve a crime. The famous Michelson–Morley experiment (see Amit Roys' article in this issue), is also a case of a null result – fringes which did not move.

From *Box* 2, the baseline at which the fringes first disappear for a uniformly illuminated disc of angular diameter Θ is given by $b_{\max} = 1.22\lambda/\Theta$. For a wavelength of 550 nm and $b_{\max} = 3.07$ m, this gives $\Theta = 2.2 \times 10^{-7}$ radians = 45 milliarcseconds.

Multiplying the intensity as a function of angle by a trigonometric function depending on the baseline is nothing but a Fourier transform.

Emboldened by this success and wanting to target other stars, whose diameters were smaller, Michelson ventured to build a 15 m beam. However, this was not successful, and nor did anyone else match this feat for decades. This brings out both the difficulty of the technique and the skill of the experimenters.

5. Aftermath

Michelson himself generalised the expression for fringe visibility to objects more complicated than uniformly illuminated circular discs. Multiplying the intensity as a function of angle by a trigonometric function depending on the baseline is nothing but a Fourier transform. But characteristically, he did not say this. So it was left to van Cittert and Zernike (both from Netherlands), in the 1930s, to bring out the general idea that one could relate the ability of two parts of the incoming light to interfere (fringe visibility or coherence) to the source properties by such a transform. In the 1950s, Emil Wolf, then in England, made the very important theoretical point that most optical measurements are really measurements of coherence, rather than electric fields. Rewriting optics in this way brings it closer to experimentally measured quantities, rather than the underlying electric fields. Also, this reformulation goes over more smoothly into quantum optics, a very significant step taken by George Sudarshan[4].

[4]Urjit A Yajnik, Symmetry and Mathematics, *Resonance* Vol.20, No.3, pp. 264–276, 2015; The Conception of Photons - Part 1, *Resonance*, Vol.20, No.12, pp.1085–1110, 2015; Part 2, Vol.21, No.1, pp.49–69, 2016.

On the astronomical side, this relation remained dormant till it was brilliantly used by another community, working with radio waves. Many people contributed to the application of Michelson's interferometer concept to radio waves, but undoubtedly the dominant figure was Martin Ryle of Cambridge University, England, who received the Nobel Prize in 1974. He was able to use pairs of radio telescopes which could be moved (and which were in any case rotated by the Earth!) to map out the full visibility function *Figure* B, and make maps of complex distributions

Figure 4. A view of six antennas (out of a total of 30) of the Giant Metrewave Radio Telescope in Khodad, near the Pune Nasik road. The waves falling on each are converted into signals on optical fibres, then taken to a central building where they are 'interfered' in all possible ways, making for 435 Michelson interferometers operating continuously as the Earth rotates, giving a good coverage of baselines varying in both length and direction. Image Courtesy: National Centre for Radio Astrophysics (www.ncra.tifr.res.in).

of radio waves emitted by astronomical objects. I cannot resist including *Figure* 4, which is the Indian effort in this direction. The picture shows some of the antennas of the GMRT (Giant Metrewave Radio Telescope) of NCRA-TIFR (National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research). This array of 30 telescopes is located near Pune, and can be viewed as $30 \times 29/2 = 435$ Michelson interferometers, operating day and night, round the year, to explore the universe with radio waves.

After 1980, Michelson's principle, with many improvements, was again applied to stars using visible light. The availability of electronic control systems and CCD detectors made much more automation and accuracy possible. Today, there are many such projects, going upto a few 100 m in baseline, gazing at the stars from many parts of the world, giving unique information on their binary nature, sizes, shapes distorted by rotation, and even surface features like spots, variability. For clarity, this article has treated the light as perfectly monochromatic. In practice, a filter can be used, or an arrangement like a prism to separate the fringe patterns at different wavelengths. In fact, this yields even more information about the star. The current status is covered in the reference given in Suggested Reading.

Suggested Reading

[1] A A Michelson and F G Pease, *Astrophysical Journal*, Vol.53, p.249, 1921, available on the web at: http://adsabs.harvard.edu/abs/1921ApJ....53..249

[2] Gerard van Belle gives a nice account of the history upto modern times in his talk, available at: http://nexsci.caltech.edu/workshop/2003/2003_MSS/07_Monday/history_030706a.pdf (This includes the original French of Fizeaus proposal)

Address for Correspondence
Rajaram Nityananda
School of Liberal Studies
Azim Premji University
PES South Campus
Electronic City
Bengaluru 560 100, India.
Email: rajaram.nityananda@gmail.com