

Basics of Optical Interferometry: A Gentle Introduction

Gerard T. van Belle

Lowell Observatory
email: gerard@lowell.edu

Abstract. The basic concepts of long-baseline optical interferometry are presented herein.

Keywords. stars: fundamental parameters, techniques: interferometers, techniques: high angular resolution, etc.

1. Introduction: Interferometry is Inevitable

Stars are, from an angular size standpoint, small. This rather qualitative, relative statement can actually be easily quantified in a back-of-the-envelope fashion. Using our sun – an object of 30 minutes of arc in angular diameter[†] – as a prototype, we can rapidly derive the size regime in which we must work to be directly examining the sizes, shapes, and ultimately surface morphologies of stars. Since the sun delivers a (mildly staggering) apparent magnitude of $V \approx -26$, with the next nearest star (in terms of brightness) at about $V \approx 0$, we see that from $V_{\odot} - V_{\text{star}} = -2.5 \log(I_{\odot}/I_{\text{star}})$ that the difference in intensity is roughly a factor of 10 billion. Under this back-of-the-envelope approach, with all stars having identical surface brightnesses, intensity scales simply with disk area, and we estimate angular diameter θ :

$$\frac{I_{\odot}}{I_{\text{star}}} = \frac{A_{\odot}}{A_{\text{star}}} = \left(\frac{\theta_{\odot}}{\theta_{\text{star}}} \right)^2 \quad (1.1)$$

From our value of 30' for the sun, we arrive at a size of 12 milliarcseconds (mas) for the nearest, brightest stars; this number diminishes rapidly into the sub-mas regime for sample sizes greater than one or two dozen.[‡]

Given that conventional ground-based telescopes are limited by the atmosphere to resolutions of roughly 0.25-0.50 arcseconds (under ideal observing conditions that occur less than 10% of the time), already this regime is out of reach. Adaptive optics ameliorates the atmospheric resolution limit somewhat, but aperture size still imposes its limits - eg. the Keck 10-m telescopes can reach a limiting resolution of $1.22\lambda/D \approx 30$ mas (assuming AO-enabled *J*-band observations). Again, an angular resolution that is insufficient for the task of stellar surface imaging.

This problem becomes even more daunting when considering things we may wish to accomplish in the not-so-distant future. The notion of surface imaging of extrasolar planets is one that provokes the ‘giggle factor’ response in many, and yet the ravenous hunger of the field for increased knowledge of these objects will ultimately lead to considering solutions for this challenge. A 10,000km object at a distance of 10 parsecs is roughly 10

[†] About the size of your thumbnail at arm’s length.

[‡] As an aside, much of the concepts in this article are presented in far greater detail – and far more precisely – in the summer school writeups such as that of Lawson (2000).

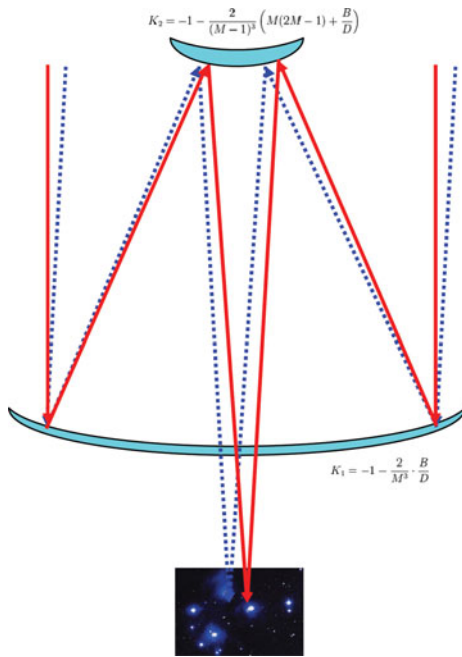


Figure 1. Path lengths through a ‘conventional’ single aperture telescope. All paths for the on-axis star (solid lines) from the star, through the atmosphere, and through the optical system, to the detector, are equal to optical tolerances ($\sigma_d < 100\text{nm}$); the same is true of the off-axis star (dotted line). Although only the edge light rays from both stars are shown, this applies to the entire ensemble of rays across the entrance aperture. Meeting this pathlength equality condition dictates the specific shapes of the primary and secondary mirrors, relative to the sky and relative to each other.

microarcseconds (μs) in size – a regime where clearly the highest angular resolution capabilities will be needed. The need for angular resolution at this scale is intimidating but any future astrophysics roadmap ultimately arrives there: *interferometry is inevitable*.

2. Interferometry is Already Here

Although you, gentle reader, are now faced with the prospect of interferometry being inevitable, do not despair! For one main unrealized lesson of many is that **all telescopes are interferometers** – the technique of interferometry is already here throughout astronomy. Illustrated in Fig. 1 is this little-acknowledged fact, which can be best thought of in this context as a question of path length: all paths through the instrument, from stars in the field of view to detector, must be equal to optical tolerances. If this condition is not met, the telescope will not be diffraction limited.

The best example of failing to meet this condition is the original Hubble Space Telescope: the shape of the primary, rather than meeting its prescription to its specified tolerance of roughly $\lambda/100$, actually came in at $10 \times \lambda$. However, since the error in the primary was a smooth deviation, corrective optics could be introduced into its optical path to return the overall system to a condition where the pathlength equality condition was satisfied[†].

A second instructive example here are the largest telescopes currently available to astronomy – those based upon giant segmented mirrors, such as the existing Keck-1 and Keck-2 telescopes, or the planned E-ELT, TMT, or GMT telescopes. These facilities have primary mirrors (and in the case of the next generation, secondary mirrors as well) that are built up of many individual elements. As a whole, the general shape of the resulting mirrors still satisfy the pathlength equality condition, albeit with slight gaps between the individual segments. If one were to increase the size of the gaps between the segments, while maintaining the pathlength equality condition, a useful diffraction-limited signal

[†] You can think of it as an adaptive optics system with an actuator response rate of 10 nHz.

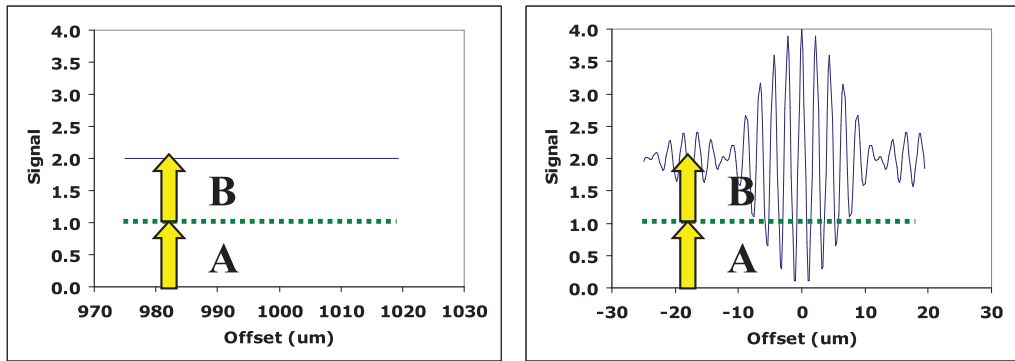


Figure 2. Photometric signal as described in §3 for a two telescope interferometer observing a point source, combined without matching pathlengths (left), and as the zero-path condition is met (right). The peaked appearance of the right-hand signal leads to the moniker ‘fringes’.

would still fall upon the detector. In this fashion one could leverage the benefits of such a telescope being an interference device – an interferometer – while increasing its resolving power.

In the end, a long-baseline optical interferometer (LBOI) is simply such a device: the space between its mirrors segments is large – very large – but it is still meeting this important pathlength equality condition. The additional wrinkle introduced with LBOI is that the segments have been replaced by telescopes themselves; the practical impact of this wrinkle is that the on-sky field-of-view of a LBOI is much smaller than that for a conventional telescope.

3. What Does a Simple Interferometer ‘See’?

For the simplest case of a two-telescope interferometer, one may think of the light collected from the two telescopes – let’s call them ‘A’ and ‘B’ – passing through the optical system, meeting up with each other at a 50% transmission / 50% reflection beam splitter, and one of those two combined beams then being concentrated upon a single-element photodetector†.

If the light from ‘A’ and ‘B’ arrives at the photodetector without the paths through each of the interferometer’s arm, through the atmosphere, and back to the star, being exactly equal‡, all that the detector will register is the amount of light from telescope ‘A’ in combination with ‘B’ (Fig. 2, left). However, if care is taken to match the pathlengths from the two telescopes, a much different signal is seen – the signal will appear to fluctuate above and below the A+B level (Fig. 2, right). This is the phenomenon of constructive and destructive interference happening as a zero-path differential is achieved.¶

If our maximum constructive (V_+) and destructive (V_-) signals are measured, we can characterize the amplitude of these ‘fringes’ simply as:

$$V = \frac{V_+ - V_-}{V_+ + V_-} \quad (3.1)$$

† This is in fact the *exact* architecture for some of the early facilities and/or ‘first light’ instrumentation of later facilities: IRMA (Dyck *et al.* 1993), IOTA (Dyck *et al.* 1995), CHARA (ten Brummelaar *et al.* 2005).

‡ Well, ‘exactly’ here means to within a fraction of operational wavelength.

¶ A frequently asked question: “Is this ‘creating’ signal on the detector?” No. Bear in mind that the 50/50 beam splitter has *two* outputs; the other of the two combined beams shows an equal but opposite signal from Figure 2, right – conservation of energy applies even here.

which is often referred to as the **visibility**. For a perfect optical system (including an atmosphere that does not degrade the signal) observing an unresolved star, $V = 1$.

This is all good and well, but where things get interesting is with a resolved star (Figure 3). A resolved star can be thought of as sending light through one's optical system from each of two halves. Each side of the star individually is unresolved, but the angular separation upon the star corresponds to a slightly different zero-path condition through the optical system. Since the interference patterns of each half is seen simultaneously, they smear each other out, reducing the contrast or viability of the star's signal. The reason why this is interesting is that there is a direct relationship between the amount of reduction in visibility and the angular size of a star:

$$V = \frac{2J_1(x)}{x}, \text{ where } x = \frac{\theta_{\text{UD}} \pi B}{\lambda} \quad (3.2)$$

with J_1 is the first-order Bessel function, θ_{UD} is the angular diameter of an equivalent uniformly illuminated disk[†], B is the *projected* baseline between the two telescopes, π is simply the familiar numerical constant, and λ is the operational wavelength.

As such, in the most straightforward case[‡] a measurement of V can be used to directly establish the angular diameter of a star. In practice, rigorously measuring V is complicated by an uncooperative atmosphere, necessitating short integration times (typically shorter than an atmospheric coherence time, \sim a few milliseconds), and non-perfect optical systems, necessitating unresolved calibrator-science target-calibrator observational interleaves to explicitly measure the true point-response of the instrument.[¶]

The underlying physical description of what is being observed is provided by the **van Cittert-Zernike theorem** (van Cittert 1934; Zernike 1938); for a LBOI, the effective implication of this theorem is that the visibility measurements we make in this context correspond to the *real* components of the Fourier transform of the image of the object being observed.

4. Fundamental Stellar Parameters from Interferometry

Angular sizes of stars are, by themselves, not terribly useful. It is in conjunction with *ancillary* data products where such data reveal their true strength. The most immediate fundamental stellar parameter – and the necessary ancillary data product – is fairly obvious, namely **linear size**. If we know the distance to a star, the linear size is realized immediately from $R = \pi\theta$, where R is the linear radius^{||}, and π is the parallax. In many cases, the real challenge is determination of π – not always readily accomplished, even in this age of *Hipparcos* and *Gaia*.

The second immediate fundamental parameter, **effective temperature**, comes from the *definition* of luminosity, L :

$$L = 4\pi\sigma R^2 T_{\text{eff}}^2 \rightarrow T_{\text{eff}} \propto \left(\frac{F_{\text{bol}}}{\theta^2} \right)^{1/4} \quad (4.1)$$

when we divide by distance on both sides: L becomes the bolometric flux, F_{bol} , and R

[†] A poor but ultimately serviceable zeroth-order approximation of the intensity distribution across the face of a star.

[‡] The star isn't *over*-resolved, is a reasonably uniform disk, isn't non-spherical, etc. etc.

[¶] One interesting aspect of measuring V , typically at the 1-5% level, to arrive at θ in Equation 3.2: the resolution limit is in this specific regime is not $1.22 \lambda/D$ but more like $0.25\lambda/D$. Hence the spatial resolution of LBOI is typically better than people naïvely expect.

^{||} Historically, angular sizes have always been in terms of 'angular diameter', whereas linear sizes have been in terms of 'linear radius'. Another one of the many idiosyncracies of astronomy.

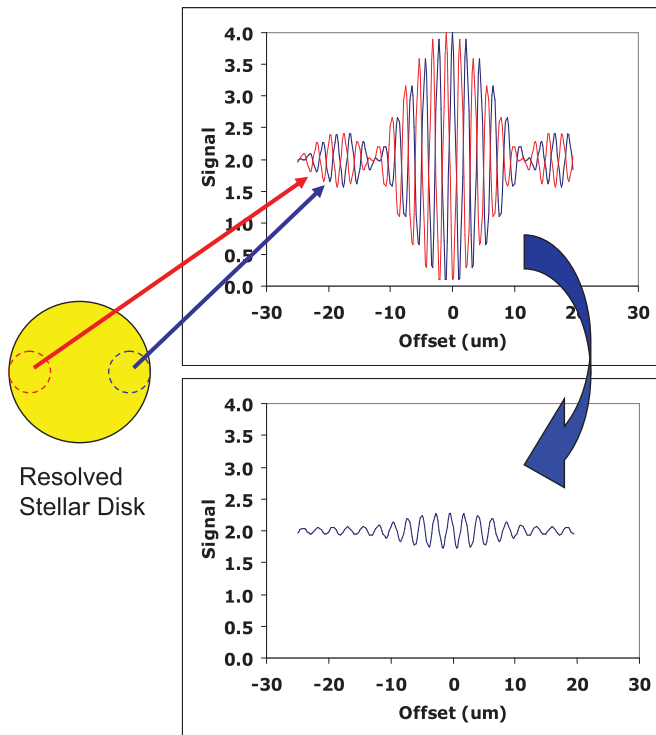


Figure 3. An optical interferometer has sufficient resolving power that, for each of the two halves of a star that is sufficiently large, there is a slight path length difference. Each half creates ‘perfect’ fringes, but the two halves combine incoherently, effectively smearing the resulting interference signal and reducing the amplitude of the fringes.

becomes θ . Here we see one of the major strengths of this approach: determination of effective temperature – an essentially *macroscopic* quantity – is being accomplished by macroscopic means, in contrast to ‘microscopic’ means such as spectroscopy. As with linear size, the real challenge in many cases is not the determination of θ but the determination of F_{BOL} . Direct observation of the entire flux budget arriving here at Earth for a star is complicated by unobservable windows, and estimates of losses due to interstellar reddening are typically somewhat model dependent.

5. What Does a More Complicated Interferometer ‘See’?

Modern facilities such as VLTI, CHARA and NPOI are building upon these fundamentals by incorporating more than 2 telescopes feeding light to the back end instrumentation – for the former, up to 4 beams can be combined simultaneously, while for the latter two facilities, up to 6 beams can be combined at once. Each pair of telescopes can in theory produce a single visibility measurement of the object being observed, with rapid gains as the number of telescopes N_T is increased – the number of visibility points N_V goes as

$$N_V = \frac{N_T(N_T - 1)}{2} \quad (5.1)$$

However, where a 3-plus element interferometer gains is in its ability to observe the **closure phase** of an object. As illustrated in Figure 4, consider the case where 3 telescope are each pointed at the same object. The true path length back to the star for a given telescope, seen in Figure 4a can be represented by an absolute phase value ϕ . Returning to the formality of the van Cittert-Zernike theorem, the phase corresponds to the *imaginary* part of a component in the Fourier transform of the image of the object being observed.

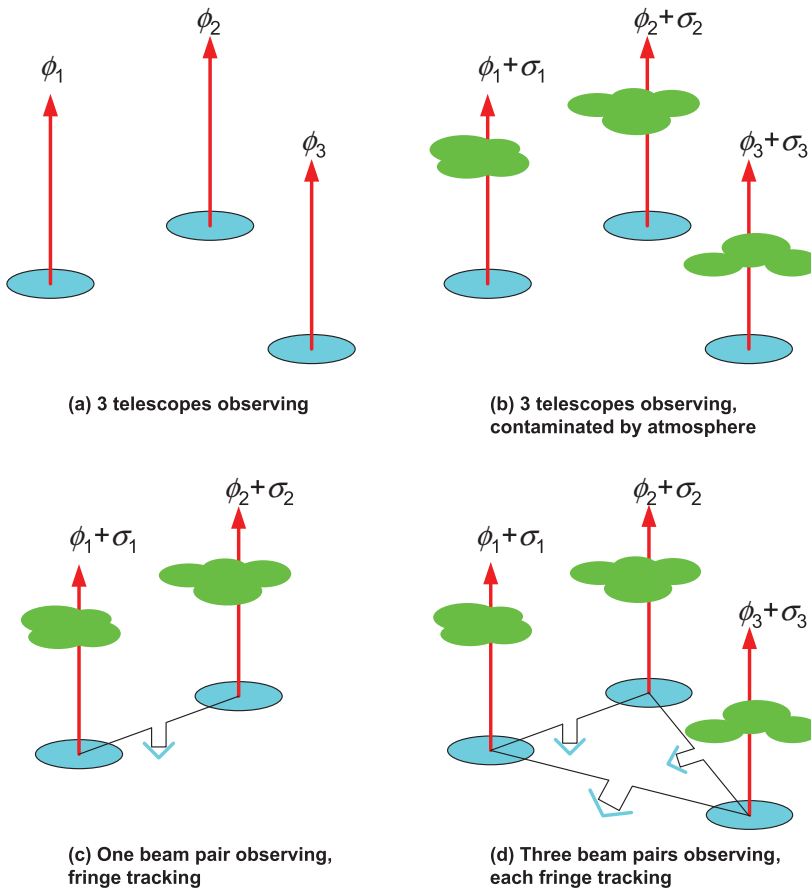


Figure 4. An illustration of observing closure phase, as discussed in §5.

Unfortunately, the atmosphere induces for each telescope an unknown amount of ‘piston’ error (Figure 4b), which is a time-variable amount of pathlength error, typically many times greater than the operational optical wavelength[†]. For even single-baseline observing with a pair of telescopes, this necessitates some form of fringe tracking, such that the position of a delay line joining the two apertures can account for the atmospheric piston error $\sigma_1 + \sigma_2$ (Figure 4c). Unfortunately, in the single-baseline case, positive acquisition and tracking of interference fringes between the two telescopes does not provide any information on the phases - the single observable, that of the necessary delay line position (Φ_{12}) to obtain fringes - is insufficient to solve for the two unknowns σ_1 and σ_2 and arrive at some knowledge of the object’s phases $\{\phi_1, \phi_2\}$.

However, in the case of 3 telescopes, the landscape alters in a subtle but significant way. In this situation, the 3 unknowns $\{\sigma_1, \sigma_2, \sigma_3\}$ are matched by 3 observables $\{\Phi_{12}, \Phi_{23}, \Phi_{31}\}$. Mathematically, we can see this by summing around the closure phase triangle:

$$\Phi_{12} = (\phi_1 + \sigma_1) - (\phi_2 + \sigma_2) \tag{5.2}$$

$$\Phi_{23} = (\phi_2 + \sigma_2) - (\phi_3 + \sigma_3) \tag{5.3}$$

[†] For example, for visible light or near-infrared observing, piston error is on the order of $\sim 1\text{-}10\mu\text{m}$, on $\sim 1\text{-}10\text{ms}$ timescales.

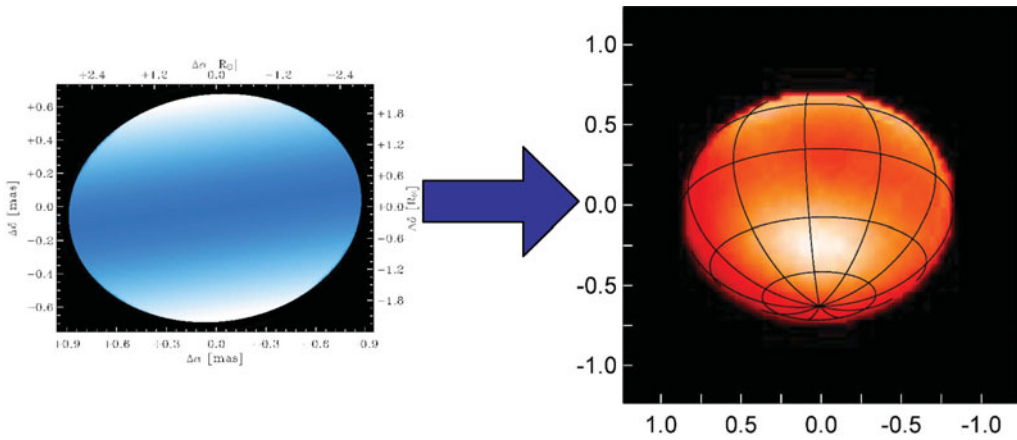


Figure 5. An illustration of the utility of closure phase: the appearance of α Cep as seen in van Belle *et al.* (2006) (left; only visibility data was available) and Zhao *et al.* (2009) (right; visibility and closure phase data). The overall scale of the object and its oblateness is well captured in both – a characteristic of visibility data – but the asymmetric surface brightness characteristic of the non-edge-on von Zeipel effect is missed in the earlier study.

$$\Phi_{31} = (\phi_3 + \sigma_3) - (\phi_1 + \sigma_1) \quad (5.4)$$

which leads to

$$\Phi_{123} = \Phi_{12} + \Phi_{23} + \Phi_{31} = \phi_1 + \phi_2 + \phi_3 \quad (5.5)$$

Although the closure phase Φ_{123} is not an individual absolute phase measurement (such as ϕ_1), we see in Equation 5.5 that it is related to the combined true phase information of the observed object and is free of atmospheric noise. In the case of a single triple of telescopes, 1/3 of the ‘true’ image information is thus recovered. As with visibility points, the number of closure phases rapidly increases with number of telescopes:

$$N_{CP} = \frac{(N_T - 2)}{N_T} \quad (5.6)$$

In the case of 4 telescopes, 50% of the object phase information is recovered; for 6, that number is 66%.

In general terms, the closure phase information is related to the imaginary part of the Fourier transform of the image upon the sky; the visibility is the real part. A more practical short-hand way of thinking of closure phase is that it is related to the degree of asymmetry of the object’s intensity distribution upon the sky. Without closure phase, significant morphology features can be difficult to recover from interferometric data (Figure 5).

6. Summary

The new 4- and 6-way closure phase instruments that are available or soon to be deployed are enabling a second revolution in optical interferometry’s impact upon astronomy through high resolution (Monnier 2007). This includes MIRC on CHARA†; MATISSE, GRAVITY and PIONIER on VLTI; and Classic and VISION on NPOI.

Overall, the technique of optical interferometry is one that has achieved a significant degree of maturity since the simple beginnings of its modern era, roughly 25 years ago.

† VEGA and PAVO are also notable new CHARA instruments but are only 3T devices.

The modern instruments and the supporting infrastructure of data analysis tools are significantly extending its community accessibility and scientific reach. Although the technique is viewed by some as still somewhat esoteric, grappling with the basic concepts of fringe **visibility** (essentially, the *size* of a interference signal) and fringe **phase** (essentially, the *position* of a interference signal) help de-mystify the underlying concepts.

References

- Boyajian, T. S., van Belle, G., & von Braun, K. 2014, *AJ* 147, 47
 Dyck, H. M., Benson, J. A., Carleton, N. P., *et al.* 1995, *AJ* 109, 378
 Dyck, H. M., Benson, J. A., & Ridgway, S. T. 1993, *PASP* 105, 610
 Lawson, P. R. (ed.) 2000, *Principles of Long Baseline Stellar Interferometry*
 Monnier, J. D. 2007, *New Astron. Revs* 51, 604
 ten Brummelaar, T. A., McAlister, H. A., Ridgway, S. T., *et al.* 2005, *ApJ* 628, 453
 van Belle, G. T. 1999, *PASP* 111, 1515
 van Belle, G. T., Ciardi, D. R., ten Brummelaar, T., *et al.* 2006, *ApJ* 637, 494
 van Belle, G. T., Thompson, R. R., & Creech-Eakman, M. J. 2002, *AJ* 124, 1706
 van Belle, G. T. & van Belle, G. 2005, *PASP* 117, 1263
 van Cittert, P. H. 1934, *Physica* 1, 201
 Whitelock, P. A., van Leeuwen, F., & Feast, M. W. 1997, in R. M. Bonnet, E. Høg, P. L. Bernacca, L. Emiliani, A. Blaauw, C. Turon, J. Kovalevsky, L. Lindgren, H. Hassan, M. Bouffard, B. Strim, D. Heger, M. A. C. Perryman, & L. Woltjer (eds.), *Hipparcos - Venice '97*, Vol. 402 of *ESA Special Publication*, pp 213–218
 Zernike, F. 1938, *Physica* 5, 785
 Zhao, M., Monnier, J. D., Pedretti, E., *et al.* 2009, *ApJ* 701, 209

Discussion

AERTS: We made several efforts to get interferometric radii of B-type pulsators (β Crucis, σ Scorpii, β Centauri) but in each case we faced the problem of systematic uncertainties due to lack of good enough calibrators resulting in highly inaccurate estimates (incompatible with other data from spectroscopy and asteroseismology). How can this be solved or improved? Can you do, e.g., 12 Lac, HD 180642 and try to get the radius of these bright B-type pulsators as test case?

VAN BELLE: Using the crude $\{V, K\}$ angular size predictor in van Belle (1999) (updated nicely in Boyajian *et al.* 2014), we can examine the rough angular sizes we should expect for these stars, which reveals some of the answer:

Star	Spectral Type	V (mag)	K (mag)	theta (mas)
β Cru	B1 IV	1.25	1.99	1.05
σ Sco	B1 III	2.89	2.40	1.15
β Cen	B1 IV	0.60	1.28	1.5
12 Lac	B1 III	5.23	5.62	0.22
HD 180642	B1 II	8.29	7.79	0.10

The key here is that these stars are *small* - angular sizes measures of less than 1.5mas are very hard to do and require facilities with the highest amounts of angular resolution; since these are all southern hemisphere objects I presume that the attempts have been made with VLTI, which would find these angular scales challenging. For the first three, a facility such as CHARA or NPOI should be able to resolve the northern hemisphere counterparts of these objects (and a corresponding set of sufficiently small calibration

objects, as described in a suitable reference such as van Belle & van Belle 2005). NPOI has greater spatial resolution by virtue of the fact it operates at $\sim 3\times$ shorter wavelengths on similar baselines to VLTI; CHARA has $\sim 3\times$ longer baselines (and is starting to operate at $\sim 3\times$ shorter wavelengths for additional spatial resolution).

The latter two stars are particularly small, but more significantly, they are very faint relative to the capabilities of modern facilities which is an additional challenge.

MASON: Regarding the relative magnitude limit of Gaia and interferometers – will the fundamental limit on diameters (eg. nearby stars) be set by Hipparcos parallax errors?

VAN BELLE: It's an interesting question where the 'fundamental limit' may be connected not so much to specific missions, but to basic astrophysics or particular objects. Cool evolved stars tend to have a certain intrinsic luminosity (eg. very roughly, $10^4 - 10^5 L_{\odot}$) which means that at a given distance, they'll have a given angular size which 'competes' with its parallax. What this translates to in slightly more quantitative terms of a specific example is that, for Mira variable stars, the angular size for a given star is typically $3\times$ greater than the parallax, so to make a parallax measurement, one is trying to measure substellar diameter photocenter shifts. Coupling this fact to the expectation that these objects will have surface morphologies and rotational periods that both change on month to year time scales – which matches the cadence of any parallax program – and you can see that there are challenges in obtaining a parallax for these objects that are above and beyond objects with smaller angular sizes (this is discussed in more detail in §3.4 of van Belle *et al.* 2002). This is a mission independent phenomenon. A sufficiently large ensemble of data can lead to meaningful results (Whitelock *et al.* 1997) but this does mean that, for individual stars, distances remain an unsolved problem.

For the hot stars of interest for this symposium, though, the relationship between their angular sizes and parallaxes mean that sub-photocenter measures are not being attempted, and this specific problem is not a concern.

On the point of Gaia: Gaia is going to be not simply an amazing mission but a truly revolutionary one; however, the specific performance in the regime of its bright limit remains to be seen. It is possible that its distances for (the relatively bright) stars of interest to interferometry will not improve substantially upon the Hipparcos results.



Gerard van Belle