The effective wavelength for broadband fringe detection using fringe scanning

W. J. Tango

30 March 2007

1. Introduction

In an interferometer like SUSI the optical path difference (OPD) z is swept around the point $z = 0$. The fringe signal appears as a "wave packet" centered on $z = 0$ with a characteristic envelope function. The area under this envelope function is proportional to the fringe visibility.

In the case of quasimonochromatic radiation the observed fringe signal is

$$s(z) = S_\sigma N_\sigma \eta_\sigma |\gamma(\sigma)| \cos[2\pi z\sigma + \alpha(\sigma)] \tag{1.1}$$

where σ is the "spectroscopic wavenumber" ($\sigma = \lambda^{-1} = c^{-1}\nu$).[1] The other quantities in this equation are:

- S_σ The overall spectral response of the optical system, including atmospheric transmission, transmission through the optics, detective quantum efficiency, etc.
- N_σ The spectral flux from the star, integrated over the apparent disk.
- η_σ A visibility loss factor that accounts for the degradation of the fringe visibility due to atmospheric effects, instrumental aberration, etc. This is strictly a stochastic quantity and in a more careful analysis should be treated as such; here I assume that we can use the mean loss factor.
- $\gamma(\sigma)$ The complex degree of coherence of the star.
- $\alpha(\sigma)$ The phase of the complex degree of coherence.

For convenience define

$$I(\sigma) = S_\sigma N_\sigma \eta_\sigma \tag{1.2}$$

for $\sigma \geq 0$ and $I(\sigma) = 0$ for $\sigma < 0$.

[1] The spectroscopic wavenumber is a proxy for the optical frequency, and avoids the necessity of writing factors of c everywhere. For optical and infrared radiation the wavenumber is conveniently expressed in units of inverse micrometers, thus a wavelength of 500 nm corresponds to 2.00 μm^{-1}.
The *circular* or *angular* wavenumber $k = 2\pi/\lambda = 2\pi\sigma$ is often used, particularly in the context of the wave equation. Like the "ordinary" frequency (f or ν) and the circular frequency ω it is easy to confuse the two kinds of wavenumber!

The goal of stellar interferometry is to derive angular information about the source from the observed fringe visibility. Indeed, according to the van Cittert-Zernike theorem the complex degree of coherence (essentially the normalized complex fringe signal) is the Fourier transform of the intensity distribution of the source on the sky. The transform depends explicitly on the wavenumber, so it must be known in order to obtain information about the source.

It is often convenient to use the "analytic signal"

$$v(z,\sigma) = I(\sigma)\gamma(\sigma)e^{-2\pi i\sigma z} \tag{1.3}$$

The actual signal can always be recovered by taking the real part of the corresponding analytic signal.

In the case of quasimonochromatic radiation there is no difficulty. However, when broadband light is used there will be a range of wavenumbers (or wavelengths) and the question arises, how do we define the *effective wavenumber*?

This turns out to be a rather subtle question, as there are several different ways of defining an effective wavenumber. Three alternative definitions are discussed in the following notes.

2. The conventional definition of the effective wavelength for broadband radiation

We first consider the rather general problem of determining the effective wavelength of a wave packet. When a wavepacket is observed with an interferometer having no dispersion it is immediately obvious that the fringes have a more or less well defined period or wavelength. If the bandwidth is wide the envelope of the fringe packet will be small and only a few fringes will be seen. Conversely, if the bandwidth is small the fringes will extend over a considerable range of optical path difference. If $\Delta\lambda$ is the bandwidth and λ the wavelength (I am deliberately not being too precise here) the width in delay space is approximately $\lambda^2/\Delta\lambda$. This of course is the coherence length of the radiation. We look at this problem in more detail.

When a broadband detector is used the fringe signal $v(z,\sigma)$ must be integrated over the bandwidth:

$$V(z) = \int_{-\infty}^{\infty} I(\sigma)\gamma(\sigma)e^{-2\pi i\sigma z}d\sigma \tag{2.1}$$

This is obviously the Fourier transform of $I(\sigma)\gamma(\sigma)$.

First consider the case when $I(\sigma)\gamma(\sigma)$ is symmetric about some wavenumber σ_e. Using the shift theorem,

$$V(z) = e^{-2\pi i\sigma_0 z}\int_{-\infty}^{\infty} I(\sigma_e+s)\gamma(\sigma_e+s)e^{-2\pi isz}ds \tag{2.2}$$

By symmetry $I(\sigma_e+s)\gamma(\sigma_e+s)$ is an even function of s [$I(\sigma_e+s)\gamma(\sigma_e+s) = I(\sigma_e-s)\gamma(\sigma_e-s)$], and it follows that the integral in Eq. (2.2) is real. Denoting this integral by by the real function $E(z)$,

$$V(z) = e^{-2\pi i\sigma_e}E(z) \tag{2.3}$$

From this it is clear that $E(z)$ is the envelope function and the fringes will have a wavelength $\lambda_e = 1/\sigma_e$.

In general, however, $I(s+\sigma_e)\gamma(s+\sigma_e)$ will *not* be an even function of s. In this case it is not immediately obvious how one might define an effective wavenumber and envelope function.

The problem can be illustrated somewhat artificially using the example of our symmetric bandpass. It is true that

$$V(z) = e^{-2\pi i\sigma'} E'(z) \tag{2.4}$$

where σ' is *any* wavenumber (this follows from the shift theorem) and the envelope function $E'(z)$ is now complex. It is of course trivial to show that

$$E'(z) = e^{-2\pi i(\sigma_e - \sigma')} E(z) \tag{2.5}$$

but the point is that we cannot, simply by inspection, determine σ_e by pulling out a phase factor from $V(z)$. We need a procedure by which we can find the "correct" envelope function from the infinite set of all possible envelope functions.

Using the argument given by Mandel (L. Mandel, JOSA **57**, 613, 1967) the true envelope function is the one which "fluctuate[s] as slowly as possible." The envelope function can be written as

$$E(z) = P(z) + iQ(z) \tag{2.6}$$

where $P(z)$ and $Q(z)$ are real functions. Following Mandel's reasoning, $E(z)$ will be as smooth as possible when the imaginary component $Q(z)$ is as close to zero as possible.

The product $I(\sigma)\gamma(\sigma) \equiv R(\sigma)$ can be written in the form

$$R(\sigma) = S(\sigma) + A(\sigma) \tag{2.7}$$

where $S(\sigma)$ and $A(\sigma)$ are the symmetric and anti-symmetric parts of $R(\sigma)$. That is,

$$S(\sigma_e + s) = S(\sigma_e - s) \tag{2.8}$$

$$A(\sigma_e + s) = -A(\sigma_e - s) \tag{2.9}$$

These functions of course depend on our choice of σ_e. The Fourier transform of $R(\sigma)$ will be

$$\tilde{R}(z) = e^{-2\pi i(\sigma_e - \sigma')}[\tilde{S}(z) + i\tilde{A}(z)] = [P(z) + iQ(z)] \tag{2.10}$$

From this we see that $Q(z)$ and $A(\sigma)$ are a Fourier transform pair. From the power theorem, if $|A|^2$ is minimized, then $|Q|^2$ will also be minimized, so we seek a way of making $A(\sigma)$ as close to zero as possible.

Multiply both sides of Eq. (2.7) by $(\sigma - \sigma_e)$ and integrate over σ. The factor $(\sigma - \sigma_e)$ is anti-symmetric about σ_e and consequently

$$\int_0^\infty (\sigma - \sigma_e) S(\sigma) d\sigma = 0 \tag{2.11}$$

from which it follows that

$$\int_0^\infty (\sigma - \sigma_e) R(\sigma) d\sigma = \int_0^\infty (\sigma - \sigma_e) A(\sigma) d\sigma \tag{2.12}$$

and the condition

$$\int_0^\infty (\sigma - \sigma_e) R(\sigma) d\sigma = 0 \tag{2.13}$$

is equivalent to

$$\int_0^\infty (\sigma - \sigma_e) A(\sigma) d\sigma = 0 \tag{2.14}$$

thus making the imaginary part of the envelope function $E(z)$ "as small as possible." It follows from Eq. (2.13) that

$$\sigma_e = \frac{\int_0^\infty \sigma I(\sigma)\gamma(\sigma) d\sigma}{\int_0^\infty I(\sigma)\gamma(\sigma) d\sigma} \tag{2.15}$$

Eq. (2.2) can now be used to with this value of σ_e to calculate the analytic signal $V(z)$. The modulus $|V(z)|$ will be the envelope function. Eq. (2.15) is essentially identical to Mandel's Eq. (34), which defines the effective frequency ν_0 (Mandel's analysis is rigorous compared to the rather handwaving argument given here).

3. The effective wavelength for $|V|^2$ measured with a scanning interferometer

The phase and amplitude of the actual fringe signal is corrupted by atmospheric effects. When the fringes are detected by sweeping the OPD the *power spectrum* of each sweep is, however, an estimate of $I^2\gamma^2$. These power spectra are then averaged and are used to estimate the fringe visibility.

The relation between the observed squared visibility and the complex degree of coherence is given essentially by Eq. (5.3) in Ireland's thesis (the integration is incorrectly over wavelength in the thesis). With some obvious changes in notation this can be written as

$$|V(\sigma_0)|^2 = \frac{\int_0^\infty I^2(\sigma)|\gamma(\sigma)|^2 d\sigma}{\int_0^\infty I^2(\sigma) d\sigma} \tag{3.1}$$

where the quantity on the left is the wavenumber-averaged square of the fringe visibility.

Expand $|\gamma(\sigma)|^2$ in a Taylor series around the effective wavenumber σ_0:

$$|\gamma(\sigma)|^2 = |\gamma(\sigma_0)|^2 + 2\gamma(\sigma_0)\gamma'(\sigma_0)(\sigma - \sigma_0) + \cdots \tag{3.2}$$

and Eq. (3.1)

$$|V(\sigma_0)|^2 = |\gamma(\sigma_0)|^2 + 2\gamma(\sigma_0)\gamma'(\sigma_0)\frac{\int I^2(\sigma)(\sigma - \sigma_0) d\sigma}{\int I^2(\sigma) d\sigma} + \cdots \tag{3.3}$$

It immediately follows that, to first order,

$$\sigma_0 = \frac{\int I^2(\sigma)\sigma d\sigma}{\int I^2(\sigma) d\sigma} \tag{3.4}$$

The accuracy of the approximation can be estimated from the magnitude of the next term in the Taylor's series:

$$\epsilon \approx [(\gamma')^2 + \gamma\gamma'']\frac{\int I^2(\sigma)(\sigma - \sigma_0)^2 d\sigma}{\int I^2(\sigma) d\sigma} \tag{3.5}$$

and for precise work one can improve the estimate of σ_0 by using successive approximations.

Note that σ_0 is the first moment of $I^2(\sigma)$, and will differ from the conventional effective wavenumber σ_e, which is the first moment of $I(\sigma)$.

4. The first moment of the power spectrum

As mentioned above, the observed squared visibility, $|V|^2$, is found by integrating the power spectrum of the observed fringe pattern and the power spectrum is the integrand appearing in Eq. (3.1). Since the power spectrum is available in the "data pipeline" we can easily calculate its first moment:

$$\sigma_1 = \frac{\int_0^\infty I^2(\sigma)|\gamma(\sigma)|^2 \sigma d\sigma}{\int_0^\infty I^2(\sigma)|\gamma(\sigma)|^2 d\sigma} \tag{4.1}$$

This differs from the definition of σ_0 by the presence of the extra factor $|\gamma|^2$ in the two integrals.

Unlike σ_0, the first moment is an observable since it can be trivially calculated from the power spectrum in the course of the data analysis. If we can find a relationship between σ_0 and σ_1 the measured σ_1 can be used as a "sanity check" that everything is OK.[2]

If the source is unresolved $|\gamma|^2 = 1$ and it is obvious that in this case $\sigma_0 = \sigma_1$. Consequently, when observing unresolved stars σ_1 can be used to directly estimate the effective wavelength. In the case of SUSI, measurements of most stars made with the 5 meter baseline fall into this category, so 5 meter measurements of σ_1 can provide the sanity check just mentioned.

In the more general case where the object is partially resolved, we can use the expansion Eq. (3.2) to relate σ_1 to σ_0. To simplify the notation set

$$\beta_n(\sigma_0) = \frac{1}{n!}\frac{d^n}{d\sigma^n}[|\gamma(\sigma)|^2]\Big|_{\sigma=\sigma_0} \tag{4.2}$$

and

$$\sigma_1 = \frac{N}{D} \tag{4.3}$$

The numerator becomes

$$\begin{aligned} N &= \int_0^\infty I^2(\sigma)[|\gamma(\sigma_0)|^2 + \beta_1(\sigma-\sigma_0) + \cdots]\sigma d\sigma \\ &= |\gamma(\sigma_0)|^2 \int_0^\infty I^2(\sigma)\sigma d\sigma + \beta_1 \int_0^\infty I^2(\sigma)\sigma^2 d\sigma - \beta_1\sigma_0 \int_0^\infty I^2(\sigma)\sigma d\sigma + \cdots \end{aligned} \tag{4.4}$$

From Eq. (3.4)

$$\int_0^\infty I^2(\sigma)\sigma d\sigma = \sigma_0 \int_0^\infty I^2(\sigma) d\sigma \tag{4.5}$$

and

$$N = |\gamma(\sigma_0)|^2\sigma_0 \int_0^\infty I^2(\sigma)d\sigma + \beta_1 \int_0^\infty I^2(\sigma)\sigma^2 d\sigma - \beta_1\sigma_0^2 \int_0^\infty I^2(\sigma)d\sigma + \cdots \tag{4.6}$$

[2]The characteristics of an optical filter, for example, may change with time.

The denominator is

$$\begin{aligned}
D &= \int_0^\infty I^2(\sigma)[|\gamma(\sigma_0)|^2 + \beta_1(\sigma-\sigma_0) + \cdots]d\sigma \\
&= |\gamma(\sigma_0)|^2 \int_0^\infty I^2(\sigma)d\sigma + \beta_1 \int_0^\infty I^2(\sigma)\sigma d\sigma - \beta_1\sigma_0 \int_0^\infty I^2(\sigma)d\sigma + \cdots \\
&= |\gamma(\sigma_0)|^2 \int_0^\infty I^2(\sigma)d\sigma + \beta_1\sigma_0 \int_0^\infty I^2(\sigma)d\sigma - \beta_1\sigma_0 \int_0^\infty I^2(\sigma)d\sigma + \cdots \\
&= |\gamma(\sigma_0)|^2 \int_0^\infty I^2(\sigma)d\sigma + \cdots
\end{aligned} \tag{4.7}$$

and to first order,

$$\sigma_1 = \sigma_0 + \frac{\beta_1(\sigma_0)}{|\gamma(\sigma_0)|^2}\left[\frac{\int_0^\infty I^2(\sigma)\sigma^2 d\sigma}{\int_0^\infty I^2(\sigma)d\sigma} - 1\right] \tag{4.8}$$

When the gradient of $|\gamma|^2 = 0$, $\sigma_0 = \sigma_1$. This of course will be true at the maxima of $|\gamma|^2$, including the maximum at the origin. This formula breaks down when $|\gamma|^2 = 0$, but see the comments in the next section.

Given a measured value of σ_1, Eq. (4.8) can be used to find σ_0 by successive approximation.

5. A caveat and an alternative approach

The underlying assumption used to derive the effective wavenumber is

$$|V|^2 = |\gamma(\sigma_0)|^2 \tag{5.1}$$

that is, we assume that the bandwidth-averaged value of the squared visibility *has the same functional form* as the quasimonochromatic squared degree of coherence. It is easy to find examples when this assumption is not true. Suppose that the object can be modelled as a uniformly illuminated disk so

$$\gamma_{\mathrm{UD}}(\sigma) = \frac{2J_1(\pi b\sigma\theta)}{\pi b\sigma\theta} \tag{5.2}$$

where b is the baseline and θ is the uniform disk angular diameter. Now suppose that we measure $|V|^2$ around the first zero of γ using a finite bandwidth. It follows from elementary calculus that $|V|^2$ can *never* be zero. This is an example of "bandwidth smearing" and has been discussed elsewhere (Tango & Davis, MNRAS **333** 642, 2002). Bandwidth smearing becomes significant when the angular size of the source is comparable or larger than the "coherent field of view" of the interferometer.

The discussion is not just an academic one. Different models for limb darkening predict values for $|\gamma|^2$ that are virtually indistinguishable when the star is only partially resolved. The differences only become apparent when $|\gamma|^2$ is measured beyond the first zero, and this is precisely the region where bandwidth smearing is important. Indeed, one has to take great care when interpreting such data to distinguish between bandwidth effects and true limb-darkening.

Reflecting on Eq. (3.1) I realized that it is actually not necessary to define an effective wavelength at all! Let $\gamma(\sigma, \theta, b)$ be the complex degree of coherence for some model, with

θ now representing possibly a set of fitting parameters (angular diameter, limb-darkening coefficients, etc.) and b is the baseline. Define a goodness-of-fit parameter for a set of N visibility measurements:

$$\chi^2 = \sum_{i=1}^{N} \frac{1}{w_i^2} \left[|V_i|^2 - \frac{\int_0^\infty I^2(\sigma)|\gamma(\sigma,\theta,b_i)|^2 d\sigma}{\int_0^\infty I^2(\sigma) d\sigma} \right]^2 \tag{5.3}$$

where w_i is the usual statistical weighting factor. We then determine θ by minimizing χ^2 using our favorite non-linear method.

This is admittedly rather more complicated than the alternative approach, which is to minimize

$$\chi^2_{\rm simple} = \sum_{i=1}^{N} \frac{1}{w_i^2} \left[|V_i|^2 - |\gamma(\sigma_0,\theta,b_i)|^2 \right]^2 \tag{5.4}$$

The advantage of this latter method is that, once σ_0 has been determined for a range of spectral types, one simply looks up or interpolates σ_0 from a table. Indeed, following common practice γ is set to $\gamma_{\rm UD}$ and a single fitting algorithm can then be used for any star as long as it has a compact atmosphere. There is absolutely nothing wrong with this approach as long as the star is only partially resolved. As a rule-of-thumb this method should be fine as long as $|V|^2 >\sim 0.05$. When $|V|^2$ is small, or one is working beyond the first zero in the visibility function, bandwidth smearing will be important and the "exact" value of χ^2 given by Eq. (5.3) should be used. Since $I^2(\sigma)$ depends on the spectral type of the source one either has to do the fitting on a case-by-case basis or to link the fitting routine to a library of spectra.