Fringe Tracking and Group Delay Tracking Methods for MIDI

J. Meisner *
Leiden Observatory

Abstract: The tracking of the white light (central) fringe in a broadband optical/IR interferometer, allows the possibility of long coherent integrations of fringe visibility. Fringe tracking involves the determination of absolute OPD offsets (not just narrowband phase) and either real-time correction of the interferometer's delay lines, or determination of the estimated OPD time series to be applied to the interferometric data off-line to effect coherent integration of fringe visibility. Algorithms of this sort will be included in the data reduction software package being developed for MIDI, the 10 micron interferometric instrument of the VLTI.

Group-delay tracking is a technique which is somewhat more robust than true phase tracking, but supplies a cruder estimate of OPD variations. Such incoherent techniques are useful for coarse adjustments to an interferometer's delay lines, and will be part of the real-time software operating in support of the MIDI instrument. General characteristics of both estimators are compared as regards sensitivity, detection bandwidth, and behavior in response to dispersion.

1 Introduction

Let us only consider a two element astronomical interferometer receiving (partially) correlated light from a star. Each of the signals received at the two telescopes has undergone a different optical path delay (OPD). The difference between the delays affecting the two beams consists of two components: 1) The geometrical component due to the position of the object in the sky relative to the interferometer's baseline, τ_g; and 2) A stochastic term due to atmospheric turbulence, which we shall denote τ_a.

Every long-baseline interferometer requires a delay-line in order to (at least partly) compensate for the relative delay affecting the beams. The geometric delay, τ_g, is deterministic, and, in principle, can be predicted. The atmospheric term is random and can only be estimated using the starlight itself. Let us assume that τ_g is known exactly and is applied to the delay line, along with a "correction" delay τ_c which may optionally attempt to follow the atmospheric delay τ_a and/or introduce a known delay offset. An interferometer using the starlight following this (partial) correction of the differential OPD, will be sensitive to the remaining delay offset $\tau = \tau_a - \tau_c$, and may be able to estimate this quantity.

There are three distinct purposes for estimating or tracking the atmospheric delay τ_a. The first reason is in order to maintain coherence in the detection of interference. For an uncorrected offset delay τ, each optical frequency will encounter a phase shift given by $\phi = 2\pi\nu\tau$. For non-zero τ, there will be a range of phases affecting the interfered light. For an interferometric channel with an optical bandwidth $\Delta\nu$, the change in phase across that bandwidth will amount

*Send correspondence to meisner@strw.leidenuniv.nl

to $\Delta\phi = 2\pi\Delta\nu\,\tau$. For a boxcar bandpass function, the *measured* fringe visibility will be reduced by the familiar factor $sinc(\Delta\nu\,\tau)$. When $\Delta\nu\,\tau$ reaches unity, the interference signal will be completely cancelled, but even much smaller OPD offsets, if not accounted for, will lead to errors in the measurement of visibility. This is avoided by controlling the delay line to insure that $\tau \ll 1/\Delta\nu$. This level of OPD estimation is often referred to as "coherencing."

The second reason for requiring estimation or tracking of the random differential OPD applies to any instrument employing coherent integration of fringe visibility. Since an interferometer measures a quantity equal to the underlying visibility rotated by the instantaneous phase shift $\phi = 2\pi\nu\tau$, the uncorrected phase shift must either be controlled so that it remains stable during the period of a coherent integration, or it must be continually estimated, so that the data can be "derotated" (Meisner 2000) prior to being co-added. In either case, the fringe tracker must be capable of determining the phase ϕ to within about a radian, a much more stringent requirement than required for "coherencing," discussed above.

Finally, an *astrometric* interferometer has the geometric delay itself as the primary observable, and relies on the fact that the random atmospheric OPD τ_a is zero mean. Thus an estimator of differential OPD is again required.

Our current work in the area of atmospheric OPD tracking will in part be directed toward the operation of MIDI, the 10 micron instrument for the VLTI (Leinert 1998). Among its various modes of operation, MIDI will sometimes require real-time corrections to the VLTI delay-lines. Additionally, estimation of differential OPD will be required for visibility estimation using coherent integration. The OPD estimation might be derived from the 10 micron MIDI data itself (as emphasized in the following discussion), or from a separate near infrared instrument observing either light from the same object, or from a close bright star using a future dual-beam facility known as PRIMA (Quirrenbach 1998). A fringe sensing unit at 1.6 microns called FINITO (Gai 2001) is being built which may be superceded by an upgraded device as part of the eventual PRIMA facility.

2 Phase Estimation of Atmospheric Delay

Consider a set of interferometric data taken at an instant, or at least over a time period short compared to the atmospheric coherence time. Given a perfect model of everything *except* τ, the maximum likelihood estimator is obtained as the value $\hat{\tau}$ which minimizes the difference between the detected signal (which includes noise) and the signal that would have been predicted on the basis of $\tau = \hat{\tau}$. Given a linear estimator of fringe visibility $\hat{V} = f(X)$ (coherent estimators are always linear estimators) then this criterion will also generally be equivalent to the one that minimizes the difference between the estimated visibility $\hat{V}$ and the true visibility $V^{(actual)}$ (which of course is not usually known!). However we can often take advantage of the fact that in many configurations, errors in the estimate $\hat{\tau}$ will tend to only reduce the real part of the estimated $\hat{V}$, at least when we have reason to believe that its phase is zero (or more generally, will reduce the real part of $e^{-j\phi}\hat{V}$ where ϕ is the true phase angle of the source visibility). That allows us to simply create an estimator which chooses the $\hat{\tau}$ which maximizes the real part of the estimated visibility $\hat{V} = f(X, \tau)$.

Let us first apply this principle to narrowband interferometric data at an optical frequency ν. Let us assume that the underlying source visibility has zero phase, so that $V^{(actual)}$ is a real number. The effect of a phase shift due to a differential atmospheric delay of $2\pi\nu\tau$, would be indistinguishable from a phase in the underlying visibility. Thus a coherent estimator which *did not* take the atmospheric phase shift into account, would yield the following (incorrect!)

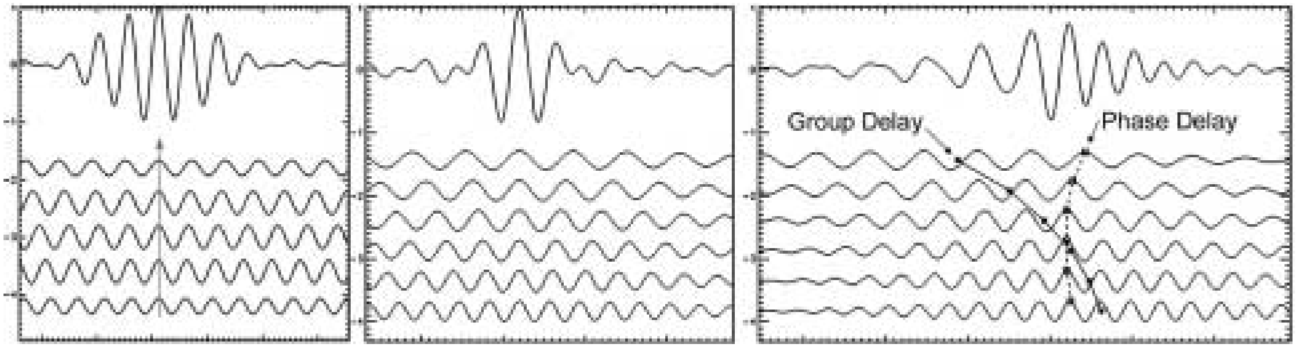

Figure 1: Left: interference at 5 different wavelengths combine to produce *global* peak of OPD likelihood as shown by arrow. Center: wideband detection (6 frequency components shown) produces pronounced peak at zero OPD, with reduced sidelobes. Right: same spectral components subject to serious longitudinal dispersion, produce "chirped" fringe with poorly defined global peak.

estimation of visibility:

$$\hat{V}_{RAW} = V^{(actual)} e^{-j2\pi\nu\tau} (+\text{noise})$$

A corrected coherent estimator, however, which took into account an *assumed* phase shift of $2\pi\nu\hat{\tau}$ would be:

$$\hat{V}_{CORR} = e^{j2\pi\nu\hat{\tau}} \hat{V}_{RAW} = V^{(actual)} e^{j2\pi\nu\hat{\tau}} e^{-j2\pi\nu\tau} (+\text{noise})$$

Again assuming that $V^{(actual)}$ is a positive real number, then the real part of $\hat{V}_{CORR}$ will clearly be maximized when $2\pi\nu\hat{\tau} - 2\pi\nu\tau = 2\pi N$ or $\hat{\tau} = \tau + N\frac{1}{\nu}$ (plus an error due to measurement noise). So this method works in principle, but only estimates τ subject to an ambiguity of an integral number of wavelengths. Of course that is a direct consequence of our narrowband model in which there *is no* distinction between delays that are separated by exactly one wavelength!

This ambiguity can be removed however, when phase information from multiple wavelengths is combined. For instance, a spectrally-dispersed interferometer might measure the amplitude of interference at five different wavelengths, producing likelihood functions each of which consists of a sine wave of this sort at a different frequency, as graphed in the lower left of Fig. 1. Each sine wave has peaks at likely positions of the OPD. However only at one point do those peaks coincide. Combining the results from interferometric detection at the five wavelengths, (in this case by simple addition) we arrive at the upper plot in which a figure of merit corresponding to a "white light fringe" has been synthesized. On the other hand such a white light fringe may have been observed directly using a delay-scanning interferometer (after filtering). In either case, the position of the unknown delay τ has been determined by the maximization of a function formed as a (weighted and/or phase shifted) sum of components produced by different wavelengths.

It may be observed that although the central peak of the combined function clearly defines the solution of the OPD estimator, the sidelobes one wavelength on either side of that peak are not so greatly reduced in magnitude. It is easy to see that the effect of detection noise could very well cause one of the sidelobes to become the global peak, leading to a misestimation of an entire wavelength. When the correct peak is observed, however, its position will be found within a small fraction of a wavelength, satisfying our criterion as a phase reference for coherent integration. This dichotomy is characteristic of phase tracking (coherent) OPD estimators: a close estimation of the actual delay accompanied by "fringe hopping" errors of one or more integral number of wavelengths. A histogram of estimation errors generated by such a simulation is shown in the left pane of Fig. 2 in which one sees frequent errors of about one wavelength (in this simulation the center wavelength was 10 microns, corresponding to 33 femtoseconds).

Integral wavelength OPD estimation errors of this sort become less frequent with a higher signal-to-noise ratio, or with data detected over a wider bandwidth. A wider bandwidth will produce lower sidelobes, as illustrated in the center pane of Fig. 1 which shows the combined fringe that would be obtained with uniform power detected over a 2:1 wavelength range. This approximately corresponds to the MIDI instrument which will be sensitive from at least 7 to 14 microns. However the right pane of Fig. 1 illustrates the deleterious effect of longitudinal dispersion that could “blur” the distinct peak seen in the center pane. Even though the same frequency components are contained in both of these plots, the shift in the phase delay of the various frequency components (plotted with small squares) causes an increasing group delay with frequency (plotted with the solid line) so that a “chirped” fringe is produced. This example illustrates the “worst case” water vapor dispersion expected to affect MIDI, corresponding to travel through a 100 meter delay line at the VLTI filled with air having 28% relative humidity at 12 degrees (Hase 2001).

In addition to dispersing the energy of the fringe in delay space, the global peak has been reduced by a quarter, barely surpassing the left sidelobe. However having calibrated the dispersion of the delay line, one can use the phase shift as a function of optical frequency to retrieve the undispersed fringe (center pane, Fig. 1) from the dispersed data. That allows for an OPD estimator (and visibility estimator) to defeat the effect of material dispersion, when known, thus increasing the estimator's ability to correctly identify the central fringe.

It should also be pointed out that these examples involve the simple addition of the data from different optical frequency components. More generally, data need to be *weighted* according to their signal and noise levels, in order to maximize the SNR of the combined figure of merit.

Beyond these corrections to data received at one point in time, a more powerful approach to the “fringe hopping” problem involves a probabilistic analysis of noisy data from a time series of interferograms. Using the *a priori* statistics of atmospheric turbulence, the smoothness of the atmospheric OPD process can help resolve $2\pi N$ phase ambiguities in the total solution from a series of interferometric data (Meisner 1996). A further description of this technique is beyond the scope of this overview, however it is applicable to all modes of interferometric data production in which a continuous data stream of sufficient SNR is produced with gaps not much larger than the atmospheric coherence time parameter.

3 Group Delay Tracking

In the literature (Nisenson 1987) “Group Delay Tracking” has referred to OPD estimation from spectrally dispersed detection, based on a somewhat different rationale, but which leads to a method not dissimilar from the phase tracking methods discussed in the previous section. We shall see that the major difference involves taking the magnitude rather than the real part of the complex likelihood function, leading to a somewhat weaker estimator, but one which is less sensitive to inaccuracies in the assumed model.

Let us recall the definitions of phase and group delay. Given phase shift as a function of frequency, the phase delay is defined as the (negative) ratio of phase to (radian) frequency, while the group delay is defined as the (negative) *derivative* of phase with respect to (radian) frequency:

$$D_P = -\left(\frac{1}{2\pi}\right)\frac{\phi}{\nu} \qquad D_G = -\left(\frac{1}{2\pi}\right)\frac{d\phi}{d\nu}$$

Since a phase, ϕ, only has meaning modulo 2π, there are multiple branches of the “phase delay” function obtained by adding $2\pi N$ to all phases. This can be seen in the right pane of Fig. 1 in which a single choice of phase peaks has been arbitrarily selected to illustrate the phase delay at 6 different frequencies. However the group delay, D_G, is unique, and can be just as well obtained from any branch of the phase delay function D_P:

$$D_G = -\left(\frac{1}{2\pi}\right)\frac{d\phi}{d\nu} = -\left(\frac{1}{2\pi}\right)\frac{d(-2\pi\nu D_P)}{d\nu} = D_P + \nu\frac{dD_P}{d\nu}$$

Although in the presence of dispersion, the group delay may not resemble the phase delay (as illustrated in Fig. 1), what is important from the standpoint of fringe tracking, is that atmospheric OPD variations are approximately achromatic, and will offset the phase delay and group delays equally and by an amount independent of optical frequency. By tracking changes in one we are just as well tracking changes in the other.

The approach taken to obtain the estimator of group delay from spectrally dispersed interferometric data, follows directly from the definition of group delay. Changes in the phase of interference between spectral channels are proportional to group delay, and it can be clearly seen that the fourier transform of a spectrum will produce a peak (in magnitude) at a group delay which corresponds to that phase difference between spectral channels. This will still be true in the case that only the real part of the correlation is available, with the consequence that the estimator produces peaks at both the positive and negative positions of the true delay, leading to the frequently stated over-generalization that the sign of the delay cannot be retrieved from such data.

Although this rationale is dissimilar from that of the phase tracking estimator (reduction of the difference between the received data and the predicted data given an assumed $\hat{\tau}$), the results are strikingly similar. Put very simply, the difference between the coherent "phase" estimator and the so-called "group delay" estimator, is that while the former identifies the OPD which maximizes the *real part* of the "synthesized fringe" or likelihood function (in the upper graphs of Fig. 1), the group delay estimator maximizes the *magnitude* of the same complex function (only the real parts of these functions have been plotted). This is equivalent to finding the peak of the envelopes plotted in Fig. 3, which are the magnitudes of the analytic signal of the delay-scanned fringe obtained from observations using VINCI.

4 Comparison of the Estimators

The difference between maximization of the real part of the fringe shown and the maximization of its magnitude, leads to qualitative changes in the types of estimation error achieved, and the tracking of the time evolution of the atmospheric OPD. As we have seen, with narrowband detection, the coherent estimator is subject to integral wavelength errors in the estimation of delay, due to noise which can confuse identification of the central fringe. However when the central fringe is correctly identified, the errors are sufficiently small (less than a radian) in order to allow the estimator's use as a phase reference for coherent integration of visibility. The group delay estimator does not have the "fringe hopping" problem at all, but has a much higher estimation error than attainable using the phase estimator (Lawson 2000). In all but very high SNR observations, its average error is in excess of a radian. Thus it is ideal for "coherencing" (keeping an interferometer's delay lines within an acceptable range) but is not suitable as a phase reference. A comparison of the errors between the two methods is shown in the histograms in Fig. 2 in which both methods have been applied to the same simulated data.

An important property of the group delay likelihood function, given by the envelope (magnitude) of the complex phase function in delay space, is that it is non-negative. This has two

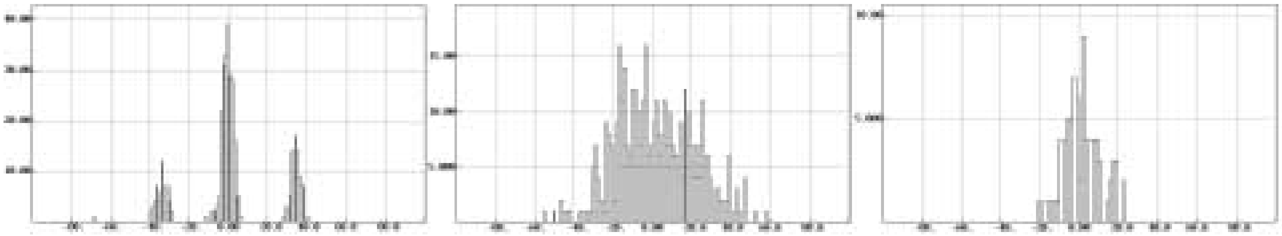

Figure 2: Histograms of OPD estimation errors using simulated narrowband10 micron interferometric data. Left: coherent (phase) estimation produces small errors with frequent "fringe jumps." Center: group delay (incoherent) estimator produces larger errors which are normally distributed. Right: group delay estimator combining 10 frames of data to improve SNR providing a consequent decrease in estimation error.

desirable consequences. First, it allows one to combine the functions from consecutive interferometric frames in order to improve the sensitivity of the estimator. It would *not* be possible to similarly combine the coherent (phase) likelihood functions which are oscillatory, for the addition of such functions produced at different times during which the OPD has varied by a wavelength would result in cancellation. Thus coherent results can only utilize data limited by the atmospheric coherence time parameter, while group delay (incoherent) results face no such restriction.

The error histogram for such a simulation, in which group delay functions from 10 frames have been combined, is shown in the right pane of Fig. 2. The average error has been reduced due to the increase in the SNR of the combined group delay functions. Combining a much greater number of samples in time would eventually lead to larger errors again, as the OPD variations during the interval of the data employed begin to dominate. However the sensitivity of the technique will increase so long as the expected change in τ does not exceed the width of the fringe envelope. Thus narrowband systems (with broad fringe envelopes, as in the left pane of Fig. 1) can better take advantage of lengthy incoherent averaging of group delay estimators in order to extract weak signals from the noise.

A second important consequence of the non-negativity of the group delay function, appears in cases where there is a substantial variation in the phase of interference over optical frequency. This could either be due to dispersion in the optical system as previously discussed, or due to the actual source structure in a well resolved astronomical source. In either case, exact knowledge of the phase function would allow us to "derotate" the raw data and recover the undispersed fringe. But lacking such *a priori* knowledge, we would find the group delay estimator to be much more robust, again, because it has no danger of adding functions of opposite polarity.

The net effect that dispersion will have, depends on the degree and type of dispersion. Especially in narrowband interferometric data, dispersion will often be of the type depicted in the right pane of Fig. 3. The two fringes shown here, and their spectra, are from two different observations by VINCI on the same night. The fringe on the left shows little dispersion, but the one on the right has a clear offset between the phase delay and the group delay, leading to a slope in the phase of its spectrum; such dispersion in the K band is almost surely due to air in the delay line (Lawson 2000). To first order, such dispersion can be described as a simple addition of a constant phase to all frequency components (which is equivalent to a slope in the phase if the arbitrary zero delay point is adjusted to constrain the phase at midband to zero as was done in these plots). For a constant phase shift θ therefore, we would have a modified fringe $x' = e^{j\theta}x$. This would have a great impact on a coherent estimator which would try to maximize the real part of x'. Its estimate would be shifted by $\theta/(2\pi\nu)$, but more importantly,

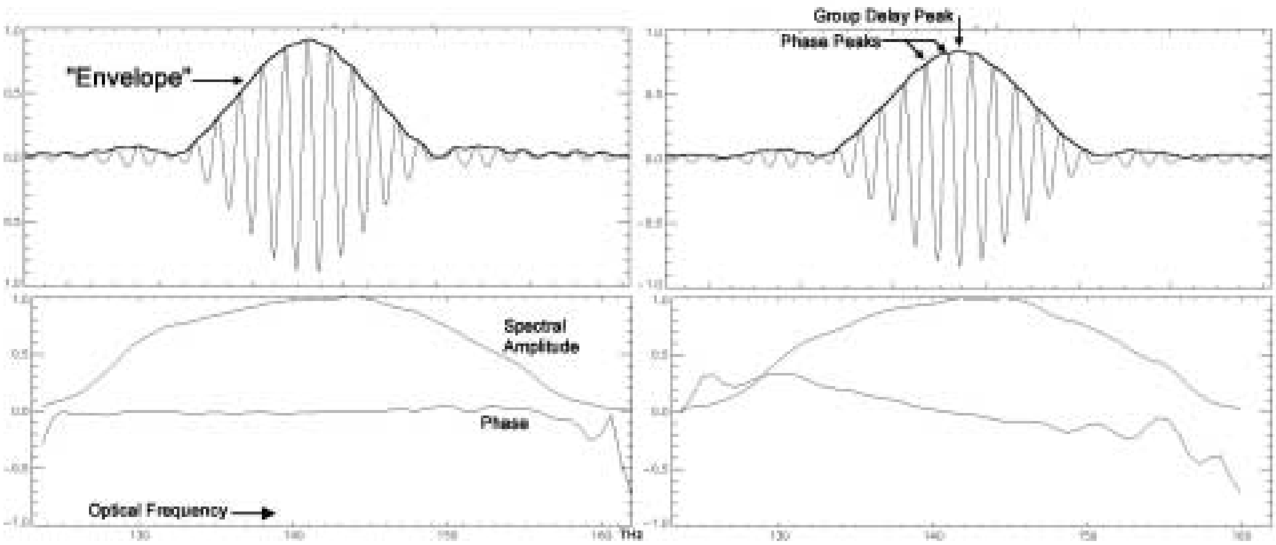

Figure 3: Average fringes measured using VINCI, the 2 micron test instrument of the VLTI. Left is from an observation of Sirius, right is V806 Centauri observed on the same night. Bottom: magnitude and phase of their fourier transforms showing change in dispersion (probably of instrumental origin). The arbitrary "zero delay" point has been adjusted to produce zero phase at midband.

the dominance of the central fringe would be weakened. This can be seen in the right pane of Fig. 3, in which there is approximately a half wavelength offset between the phase and group delays, producing two almost equal peaks of the real fringe. A naive phase estimator would be hopelessly confused!

On the other hand, the group delay estimator would see no difference between the magnitudes of x and $x' = e^{j\theta}x$. The group delay estimate is utterly insensitive to this level of dispersion! It would, however, be affected negatively by the more serious dispersion depicted in the right pane of Fig. 1. In this case, the entire envelope of the fringe has been widened. In fact, the group delay is a strong function of optical frequency (as shown), thus broadening the estimator due to the combination of these components. However even in this case, the detrimental effect of that broadening is not nearly as drastic as its effect on the phase estimator, which as previously discussed, would be confused by noise in its choice between two phase peaks of similar heights.

References

Gai M., 2001, conf. "Beyond Conventional Adaptive Optics," (ESO, in publication)

Hase F., Mathar R. J., 2001, "Water vapor dispersion in the atmospheric window around 10 microns," submitted to Appl. Opt.

Lawson P. R., Colavita M. M., Dumont P. J., Lane B. F., 2000, Proc. SPIE 4006, 397

Leinert C., Graser U., 1998, Proc. SPIE 3350, 389

Meisner J., 1996, Opt. Eng. 35, 1927

Meisner J., 2000, Proc. SPIE 4006, 1068

Nisenson P., Traub W., 1987, conf. "Interferometric Imaging in Astronomy," publ. NOAO, 129

Quirrenbach A., et. al., 1998, Proc. SPIE 3350, 807