FASR Memo

Delay Tracking, Fringe Rotation, and Phase Switching in FASR

A. R. Thompson, Nov. 26, 2007

1 Delay Tracking

For an incoming wavefront from the Sun the path lengths to different antennas of an array are generally unequal and vary as the earth rotates. The relative time differences in the wavefront arrival at the antennas are referred to as the geometric delays, τ_g. To preserve the correlation between the signals received at different antennas, it is necessary to compensate for the different geometric delays. Thus the signal received at each antenna is subjected to an instrumental delay τ_i that is adjusted continuously so that $\tau_g + \tau_i$ plus delays in transmission lines, etc. are the same for all antennas. The signals at the correlator inputs are thereby aligned in time with respect to a common wavefront, incident from the desired field center. If the instrumental delays are applied to the signals at the same frequency as that at which they are received, then the phase shifts resulting from the geometrical delays are exactly compensated by those provided by the instrumental delays. In such a case the fringe pattern tracks the field center on the sky and there are no fringe oscillations at the correlator output. In practice, however, the instrumental delays are usually introduced at an intermediate frequency, which differs from the frequency at which they are received by an LO frequency ν_{LO}. Thus the signals at the correlator input differ by a phase shift of magnitude $2\pi\nu_{LO}\tau_g$, which results in fringe oscillations at the correlator output as τ_g varies. Also, in a practical situation the adjustment of the instrumental delays is not precisely continuous but is performed in small discrete steps. These are inserted after the signals have been digitized and the sample interval τ_s provides a convenient increment for coarse delay adjustment. For Nyquist sampling, $\tau_s = 1/2\Delta\nu$ where $\Delta\nu$ is the signal bandwidth. Fine adjustment of the delay, to within a small fraction of τ_s, can be effected by introducing corrective phase shifts which vary linearly with frequency across the $\sim$ 4096 signal channels of FASR.

In an array with a large number of antennas, the instrumental delays are adjusted relative to the geometric delay for a reference antenna, which can be arbitrarily chosen, and for which the instrumental delay can remain constant. The delay error for any antenna is the difference between the total delay from the wavefront to the correlator for that antenna and for the reference antenna. When the delay error becomes as large as $\pm\tau_s/2$, the instrumental delay is adjusted by an increment $\mp\tau_s$. Thus the delay error for a single antenna is uniformly distributed over $\pm\tau_s/2$. For any pair of antennas (not including the reference antenna) it can generally be assumed that the combined delay error has a triangular probability distribution as shown in Fig. 1, with extreme values $\pm\tau_s$ and rms value $\tau_s/\sqrt{6}$. The instrumental delays are inserted after sampling, which is applied to a frequency band $N\Delta\nu$ to $(N+1)\Delta\nu$, where N is an integer. For any value of N the frequency of the sampled data is effectively within the baseband range from 0 to $\Delta\nu$. The rms frequency within this band is $\Delta\nu/\sqrt{3}$ and the corresponding rms phase error is $2\pi(\Delta\nu/\sqrt{3})(\tau_s/\sqrt{6}) = \pi/3\sqrt{2} = 42°$. For a continuum correlator the resulting loss in sensitivity is 23% (see Appendix), and results from the fact that the angles of the fringes on the sky vary with frequency, so over the full bandwidth the fringes tend to "wash out". Dividing the bandwidth into as few as, e.g., 16 channels would reduce the sensitivity loss to a few tenths of 1%. In FASR the digital signals will be filtered into $\sim$ 4096 channels before cross correlation to allow subsequent rejection of RFI. For each antenna pair the $\sim$ 4096 corresponding channel pairs will be individually cross correlated. Because of the much smaller bandwidths after channelization, the variation of the phase errors within a single channel will be negligibly small, so no loss in signal amplitude will occur: see Carlson and Dewdney (2000) for a description of a similar case. However, the delay errors cause phase errors that vary from channel to channel across the band, and these should be corrected for. At any instant the phase error is equal to $2\pi\times$ (the delay error) $\times$ (the corresponding channel frequency within the baseband range 0-$\Delta\nu$) radians. These phase corrections can be inserted in combination with the phase changes to remove the fringe oscillations considered in Section 2.

The maximum rate of change of delay occurs for a baseline of 5 km (E-W) and the Sun at zero declination,

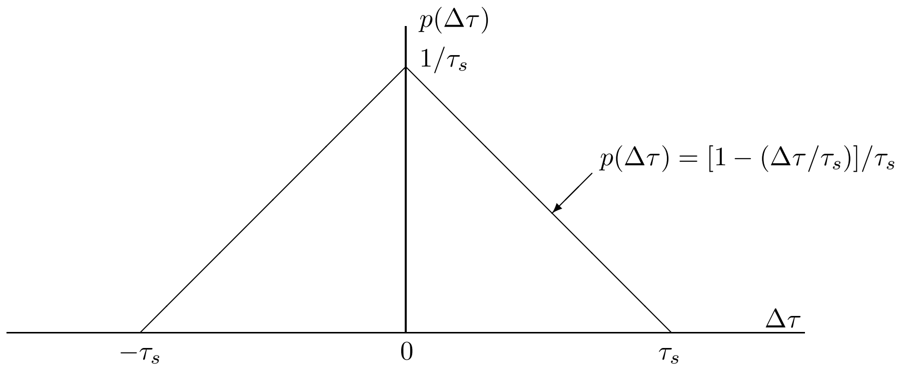

Figure 1: Probability distribution $p(\Delta\tau)$ of the delay error $\Delta\tau$ for a pair of antennas. The minimum increment of the instrumental compensating delay is equal to τ_s, the time interval between samples of the signal. The expression shown for $p(\Delta\tau)$ applies to the part of the probability function for which $\Delta\tau \geq 0$.

and is

$$\frac{d\tau}{dt} = \omega D \cos(H)/c \tag{1}$$

where D is the baseline length, H is the hour angle, $\omega = 7.27 \times 10^{-5}$ rad/s is the angular rate of rotation of the Earth relative to the Sun, and c is the speed of light. Thus, for the Sun on the meridian and $D = 5$ km, the maximum value of $\frac{d\tau}{dt}$ is 1.21×10^{-9} seconds of delay per second of time. For FASR, τ_s is chosen to be 8.33×10^{-10} s, which is the Nyquist sampling rate for a bandwidth of $\Delta\nu = 600$ MHz. (The bandwidth determined by the frequency response of the receivers is 500 MHz, which is centered within the 600 MHz.) The minimum time for the geometric delay to change by one sample interval is τ_s divided by the maximum value of $\frac{d\tau}{dt}$ which is 0.69 s. The time interval over which the correlator output is averaged has been chosen as 20 ms based on computing considerations, and since 20 ms is small compared with 0.69 s, resetting the coarse delays at fixed 20 ms intervals should not be a problem. Thus the coarse delays can remain constant during each averaging period, which simplifies the signal processing.

It is also useful to calculate the rate of change of the fine delay correction. The fine-delay phase correction is equal to $2\pi\times$ (the fine-delay error) $\times$ (the corresponding channel frequency within the baseband range $0 - \Delta\nu$) radians. The factor $\frac{d\tau_g}{dt}$ on which the change of delay error depends [see Eq. (1)] varies relatively slowly and over short intervals of time can be considered constant. The rate of change of $\frac{d\tau_g}{dt}$ is

$$\frac{d^2\tau_g}{dt^2} = \frac{-\omega D \sin(H)}{c}\frac{dH}{dt} = \frac{-\omega^2 D \sin(H)}{c}, \tag{2}$$

since $\frac{dh}{dt}$ is equal to ω which is constant with respect to time. The highest fringe frequency occurs at hour angle $H = 0$ (at the meridian), but the highest rate of change of fringe frequency occurs at $H = 90°$. For a baseline of $D = 5$ km and $H = 90°$, $\frac{d^2\tau_g}{dt^2} = 8.81 \times 10^{-14}$ seconds of delay per (second of time)^2. If the compensation for the delay is exact at time $t = 0$, then after time t the rate of change of delay is $\frac{d^2\tau_g}{dt^2}t$, the mean rate of change over time t is $\frac{1}{2}\frac{d^2\tau_g}{dt^2}t$, and the accumulated delay error is $\frac{1}{2}\frac{d^2\tau_g}{dt^2}t^2$. For the highest frequency channel, centered at $\sim$ 600 MHz, 1° of accumulated phase error occurs for $t = 10$ s. Thus the proportionality factor for the phase as a function of time, for the worst-case (5 km baseline and 600 MHz channel), should be recalculated after about 10 s.

2 Fringe Rotation

Since the instrumental delay remains constant during each of the 20 ms observing periods, the fringe frequency within these periods is equal to $\frac{d\tau}{dt}$ from Eq. (1) multiplied by the observing frequency[1]. Thus the maximum fringe frequency is $1.21 \times 10^{-9} \times 20 \times 10^{9} = 24$ Hz for 20 GHz observing frequency or 36 Hz for 30 GHz. The effect of time averaging on the fringes is to convolve the sinusoidal fringe function with a rectangular function of width equal to the averaging time. If we consider stopping fringe oscillations of frequency ν_f at the correlator output, i.e. after averaging for a time τ_{av}, the averaging reduces the fringe amplitude by a factor[2] $\mathrm{sinc}(\nu_f \tau_{av}) = \sin(\pi \nu_f \tau_{av})/\pi \nu_f \tau_{av}$. For 20 ms and 24 Hz the amplitude factor is 0.66 and for 36 Hz it is 0.34. These represent a reduction in signal-to-noise ratio for the fastest fringes, and therefore fringe stopping after 20 ms averaging is not deemed practicable. The proposal is to insert the required phase corrections into the digitized data prior to cross correlation. They are then applied to each of the N_a antennas, whereas after cross correlation the corrections would be applied to the $N_a^2/2$ antenna pairs. Subtraction of ϕ radians from the phase of a complex number $(X_r + iX_i)$ that represents a signal sample results in

$$[X_r \cos(\phi) + X_i \sin(\phi)] + i[X_i \cos(\phi) - X_r \sin(\phi)]. \tag{3}$$

It is useful to calculate the rate of change of the fringe correction, for which we need $\frac{d^2\tau_g}{dt^2}$. From Eq. (2) and the discussion following it, the maximum value of $\frac{d^2\tau_g}{dt^2}$ within FASR is 8.81×10^{-14} seconds of delay per (second of time)^2. Suppose that the fringe frequency, $\frac{d\tau_g}{dt}\nu$ where ν is the signal frequency received at the antenna, is correct at time $t = 0$. At time t the fringe frequency will have changed by $\frac{d^2\tau_g}{dt^2}\nu t$ and the accumulated fringe error will be $\frac{1}{2}\frac{d^2\tau_g}{dt^2}\nu t^2$ cycles. For baseline $D = 5$ km, $H = 90°$, and frequency 20 GHz, a fringe phase error of 1° occurs after 1.78 sec. So, in applying the fringe phase correction, the factor by which the fringe frequency is calculated should be changed within this interval in the worst (longest baseline and highest frequency) case.

3 Phase Switching

Phase switching is used in radio interferometry to mitigate errors that result from unwanted components that infiltrate the receiving channels of different antennas and introduce spurious correlation[3]. At low levels, these can be introduced through power supplies, local oscillators, monitoring circuitry, etc. Phase switching also removes the effect of any DC offset in the zero level of the digitizer. In the VLA it is found to be helpful in reducing low level unwanted responses and thus improving dynamic range. The principle involved is to modulate the phase of the wanted signal early in the system using phase shifts of magnitude π (which are equivalent to inversion of the signal voltage). These can be removed after the signal is digitized since infiltration of unwanted components does not then occur. The removal is performed by repeating the first switching sequence. Thus, at the correlator input, the unwanted components are modulated but the wanted ones are not. For each antenna a different switching function is used, any pair of which are orthogonal, resulting in elimination of the unwanted components in the averaging at the correlator output.

It is difficult to say whether phase switching is essential in observations of a strong source such as the Sun, but it is likely to be helpful in observations of calibrators. Phase switching requires a set of two-state (± 1) switching functions that are mutually orthogonal and repeat after a fixed time interval referred to as the time base. Unwanted signals are most efficiently removed by averaging for a period equal to the time base (or an integral multiple thereof). For FASR we therefore consider a time base equal to the chosen 20

[1] If the instrumental delays were varied continuously rather than in increments, the fringe frequency would be equal to $\frac{d\tau}{dt}$ multiplied by the difference between the observing frequency and the frequency at which the instrumental delays are inserted.

[2] For example, the amplitude factor takes values of 0.99, 0.98, 0.95, and 0.90 for values of $(\nu_f \tau_{av})$ 0.078, 0.1106, 0.176, and 0.250, respectively. An averaging time of 1/13 of the fringe period produces an amplitude loss of approximately 1%, which is about the maximum generally tolerable.

[3] In cases where there are many fringes within the data averaging time, as in VLBI, phase switching is not necessary. However, in FASR, for the fastest fringes ($\sim$ 24 Hz) only half a cycle occurs within 20 ms.

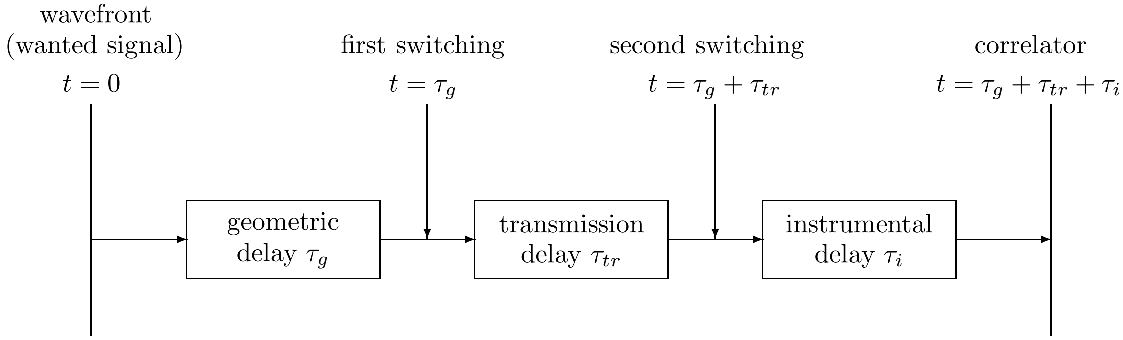

Figure 2: Delays in the FASR system that are large enough to affect the correlation of the Walsh functions used in phase switching. Unwanted components that infiltrate the system after the geometric delay but before the instrumental delay are removed by the phase switching. Here t is time relative to a signal wavefront that intercepts the delay reference antenna.

ms averaging period. Walsh functions provide convenient orthogonal switching functions: see, for example, Beauchamp (1975), Emerson (2005), or Thompson et al. (2001, pp. 242-246). Let the shortest interval between two transitions in a set of Walsh functions be Δ. The time base for the Walsh function set, which is the time interval after which all the functions repeat, is equal to Δ multiplied by a power-of-two integer, n. For any chosen values of the time base and the minimum interval Δ, there are n mutually orthogonal Walsh functions[4]. Thus, we can choose n as the lowest power-of-two integer that is greater than N_a which, for the nominal FASR value of $N_a \simeq 60$, is 64. We choose the time base to be equal to the required data-averaging time of 20 ms for FASR, in which case $\Delta = 20/64 = 0.312$ ms.

In designing the phase switching system details of timing must be considered. In general, the orthogonality of Walsh functions requires that there should be no relative time shifts between the functions[5]. The first switching occurs at the antennas and the second one after digitization at the central electronics facility. Between these two points the signals suffer a transmission time delay τ_{tr} in optical fiber or cable, which is constant in time but different for each antenna. The major system delays are shown in Fig. 2. Delays in the analog or digital circuitry can generally be neglected since they are small. There are three main timing requirements. (1) As discussed in Section 1, the total delay from the incident wavefront to the correlator input, $\tau_g + \tau_{tr} + \tau_i$, must be the same for all antennas. This is implemented through adjustment of τ_i to maintain the correlation of the wanted signals. (2) The second phase switching should follow the first with a delay τ_{tr} so that the transitions in the wanted signals will be aligned to cancel precisely. (3) Both switchings should be delayed by the geometric delay of the antenna, τ_g, so that at the correlator input, the switching transitions in the unwanted components are aligned in time from one antenna to another.

The delay τ_g in the two switchings described above needs to be updated periodically as the geometric delays change. It is useful to consider the effect of omitting this delay since this would somewhat simplify the timing system. This depends upon the overall timing accuracy required for the Walsh transitions. For the wanted components, consider the effect of a small time offset δ in the timing of the first and second switchings. For each transition, the cancellation of the first switching fails to occur for a period δ. In the cross correlation with the signal from another antenna, the correlator is reversed for a period δ and thereby cancels an equivalent interval of the unreversed output. Thus, for each transition, there is an effective loss

[4]For any set of n Walsh functions, $n/2$ are even with respect to the center of the time base, and $n/2$ are odd. These two sets are referred to as Cal and Sal functions, respectively, by analogy with the symmetry of cosine and sine functions. The sequency of a Walsh function is defined as twice the number of transitions within the time base. Any complete set of n Walsh functions contains one with no transitions, i.e. zero sequency.

[5]Pairs consisting of an even (Cal) an odd (Sal) functions remain orthogonal in the presence of time shifts, but such combinations are only possible for half the antenna pairs.

of signal for a period 2δ. The average fractional loss of sensitivity is $2n_t\delta/\tau_{tb}$ where τ_{tb} is the time base, and n_t is the number of transitions within the time base (i.e. half the sequency). Omitting the delay τ_g in the two switchings would result in an offset equal to the geometric delay. For a baseline of 5 km the maximum geometric delay is 16.67 μs. The equivalent signal loss is 33.3 μs per Walsh transition, which is 1/600 of the 20 ms time base. Thus if the maximum tolerable loss in sensitivity is 1%, and if the τ_g delay in the switchings is omitted, the maximum allowable number of transitions within the time base is 6. However, this is for the longest baselines, and higher numbers of transitions would be tolerable for shorter baselines. Thus, with a careful assignment of the sequencies, omission of the τ_g delays in the switching may be tolerable. A full analysis of the sensitivity loss would require details of the antenna locations. For calibration observations using radio sources with accurately known positions, which will be necessary to establish parameters of the array such as the relative positions of the antennas, the data averaging times can be extended to several seconds, i.e. at least two orders of magnitude greater that the 20 ms required for the solar observations. The intervals between Walsh transitions can thus be correspondingly increased, in which case any effect of the τ_g delay in the switching times is entirely negligible. Note that for solar observations the effect of omitting the delay in the switchings is small, but the effect of doing so varies with the position of the Sun in the sky, and would not be removed by calibration observations.

The effect of a timing offset on the rejection of the unwanted components depends on the loss in orthogonality of the Walsh functions at different antennas. This is more complicated than the effect of an offset on two identical Walsh functions discussed above. The loss in orthogonality depends upon the sequencies of the two functions involved, as shown by Emerson (2005). For example, a Walsh pair consisting of a symmetric (Cal type) and an asymmetric (Sal type) function, of the same set, remain orthogonal in the presence of a time offset. In other cases, the remnant of the correlation of the unwanted components is never more than the loss in correlation for a pair of identical functions [see Fig. 3 in Emerson (2005)]. Thus the restriction on the number of transitions n_t, discussed above for the wanted signals, also ensures that the unwanted responses are reduced by a least a factor of 100. It is clearly beneficial to use equal numbers of Cal and Sal type functions in the array so that for at least half of the antenna pairings the orthogonality is independent of time offsets.

In both the filtering of the signals into the ~ 4096 channels and the Hilbert transformation of the signals, the computation of each output sample involves a short time sequence of signal samples which should not contain a phase reversal. Therefore, the second phase switching, which removes the phase transitions from the wanted signals, should occur before the filtering and Hilbert transformation steps. The fine delay adjustment can be ignored with respect to its effect on the phase switching, and the sequence of operations that occurs after digitization of the signals should be as follows.

1.) Apply the second phase switching by applying sign reversals to the samples.
2.) Insert coarse instrumental delays.
3.) Apply the digital filtering algorithm to filter the signal into 4096 frequency channels.
4.) Decimate the data to the Nyquist rate for the channel bandwidth.
5.) Apply a Hilbert transform to provide an imaginary part so that the data are then in complex form.
6.) Adjust the phase of each sample to correct for the residual delay errors and remove the fringe oscillation.
7.) Perform the cross correlations.
8.) Apply the 20 ms averaging.

The second phase switching could be applied after insertion of the coarse delays, but in that case the timing of the switching would have to include a further delay equal to the coarse delay, which varies with time.

The ~ 60 Walsh functions required for FASR can each be stored as 64 bits at an antenna and read out at intervals $\Delta = 0.3125$ ms. To include the τ_g delay in the timing of the two switchings, it would be adequate to update the delay of the switching whenever τ_g changed by $\sim 1\mu$s. The maximum rate of change of τ_g for the 5 km baseline is $1.21 \times 10^3\,\mu$s per second (from Section 1) or 1 μs in 14 min.

4 Conclusion

The following limitations on the on the signal processing result from consideration of 20 GHz observing frequency and 5 km baselines.

1.) The coarse delays can be reset at 20 ms intervals, i.e. between signal averaging periods.

2.) For the fine delay corrections, it can be assumed that the required phases remain linear with time for up to 10 sec for the highest frequency (600 MHz) channel.

3.) The approximate maximum fringe frequency is 24 Hz for 20 GHz frequency and 36 Hz for 30 GHz.

4.) For the fringe rotation correction it can be assumed that the correction remains linear with time for up to 1.8 sec.

5.) For up to 64 antennas the minimum interval between switch transmissions for any Walsh function is 0.312 ms, and equal numbers of the Cal and Sal forms should be used. Functions with the lowest number of transitions within the time base should be used for the most distant antennas.

To determine tolerable limits on various parameters above, east-west baselines and the highest fringe frequencies, etc. have been considered. For calculation of geometric delays and fringe frequencies for any baseline and hour angle see, for example, Thompson et al. (2001), Eq. 4.1, in which the component w is equal to the path length to the antenna relative to the $X_\lambda, Y_\lambda, Z_\lambda$ coordinate origin, measured in wavelengths. If X_λ, Y_λ, and Z_λ in Eq. 4.1 are replaced by X, Y, and Z measured in meters, w becomes the path difference in meters which, divided by c, is the geometric delay relative to the coordinate origin. The fringe frequency is the first derivative of $\nu\tau_g$ with respect to time.

Appendix

To calculate the loss in sensitivity resulting from the incremental adjustment of the instrumental compensating delays we note that for a pair of antennas the combined delay error $\Delta\tau$ introduces an error in the fringe phase. For a component of the signal at frequency ν, the phase error is $2\pi\nu\Delta\tau$. The output of the correlator is proportional to the cosine of the phase error. For the component at frequency ν the variation of the delay error shown in Fig. 1 results in a response equal to

$$2\int_0^{\tau_s} p(\Delta\tau)\cos(2\pi\nu\Delta\tau)d\Delta\tau = \frac{2}{\tau_s}\int_0^{\tau_s}\left(1-\frac{\Delta\tau}{\tau_s}\right)\cos(2\pi\nu\Delta\tau)d\Delta\tau = \left[\frac{\sin(\pi\nu\tau_s)}{\pi\nu\tau_s}\right]^2. \quad (4)$$

Then the response factor for the full frequency band 0 to $\Delta\nu$ is equal to

$$\frac{1}{\Delta\nu}\int_0^{\Delta\nu}\left[\frac{\sin(\pi\nu\tau_s)}{\pi\nu\tau_s}\right]^2 d\nu = 2\int_0^{\frac{1}{2}}\left[\frac{\sin(\pi x)}{\pi x}\right]^2 dx = 0.774, \quad (5)$$

where we have put $x = \nu\tau_s$ for convenience in numerical evaluation of the integral.

References

Beauchamp, K. G., Walsh Functions and Their Applications, Academic Press, 1975

Carlson, B. R. and P. E. Dewdney, Efficient Wideband Digital Correlation, Electronics Letters, 36, 987-988, 2000.

Emerson, D.T., Walsh Function Demodulation in the Presence of Timing Errors, Leading to Signal Loss and Crosstalk, ALMA Memo. 537, NRAO, 2005.

Thompson, A. R., J. M. Moran, and G. W. Swenson, Interferometry and Synthesis in Radio Astronomy, John Wiley, 2001.