VLB ARRAY MEMO No. 345

PHASE AND DELAY TRACKING AT STATIONS: PRELIMINARY DESIGN

Larry D'Addario and Jon Romney
13 April 1984

This memo will describe a rough design for a system of signal processing in which fringe rotation is done at each station before recording. We will attempt to unify a number of "unconventional" signal processing concepts — including delay tracking and phase switching at the stations, in addition to fringe rotation -- into a concrete and coherent scheme. While these elements are logically related, they can be implemented independently, and should be accepted or rejected separately.

The concepts involved are "unconventional" only in the context of historical VLBI practice, and are in fact well-proven standard practice in connected-element interferometry, in particular at the VLA. This scheme may be more appropriate to a dedicated VLBI array. Indeed, the VLBA configuration will provide baselines sufficiently short to blur many of the distinctions between VLB and conventional interferometry.

Station Equipment

(a) Fringe Rotators

The fringe rotation cannot be applied in the first LO, because the total LO frequency on each channel is generally different, so the fringe frequencies are different. The rotation must therefore be applied to the baseband conversion LO. Notice that if both upper and lower sidebands are being used, then the commonly used trick of offsetting the fringe rotation to the middle of the band to compensate for delay errors cannot be done; fringe rotation must be at the LO frequency.

Each of the 16 LOs at each station must be phase shifted by an amount corresponding to the time taken for a wavefront from the tracking direction in the sky to travel from the station to a reference station. One could minimize the maximum fringe rate by choosing the reference station to be in the center of the array, but one does only slightly worse by choosing it at the center of the earth; we take the latter for convenience. Then, at 86 GHz, the maximum rate and acceleration of phase are 8.37×10^5 radians/sec and 60.9 radians/sec^2, respectively. An oscillator whose frequency and initial phase are programmable would need to be updated every 1.6 msec in order to track within 0.1 radian; this results in a data rate which is too high for the monitor/control bus. However, an oscillator whose rate of change of frequency is also programmable would only need updates about every 10 sec, and this is quite feasible to implement.

(b) Samplers

We assume that variable phase sampling will be implemented, as discussed in VLBA Memo No. 333. A single sampling clock will suffice for all channels.

(c) Phase Switching

In order to cancel d.c. offsets in the quantizers, and also to cancel spurious correlated signals, a phase-switching system will need to be implemented. (Such phase switching would probably be needed even without fringe rotation at the stations, unless steps are taken to ensure that the fringe rates are always high; the latter approach has its own problems.) At each station, the phase would be shifted 180 degrees exactly half of the time, on a schedule determined by a Walsh function; orthogonal Walsh functions would be used at all stations. The highest sequency function would have to be 16 times the lowest in order to accomodate 19 stations (32 orthogonal functions), so there would be a minimum integrating time of 16 times the shortest available switching period; if the latter is 50 msec (20 Hz), we must always integrate for 0.8 sec to cancel the offsets. Switching at 160 Hz would allow integrations as short as 0.1 sec.

The phase switching can be inserted in any LO. It will probably be convenient to insert it in the baseband converter LO along with the fringe rotation, but there may be technical problems with this; if so, it could easily be done in the first LO.

The phase switching could be removed by a digital sign reversal immediately after digitizing and before recording. In that case, nothing need be done at the correlator. However, it is possible that in some recording schemes the transmission errors may be subject to bias, and this can also cause a d.c. offset. To cancel this effect also, the phase switching should be removed at playback time. This can be done by including a bit in each header to specify the state of the phase switch and ensuring that the switch changes state only between frames.

(d) Accountability

It is probably not practical to log all of the commands sent to the fringe rotators during an observation, but this might be done if absolutely necessary. Instead, we suggest merely logging all geometrical and hardware parameters which were in use at observe time; the correlator or post-processing computer could then reconstruct the observe-time processing if necessary.

The accountability requirements are identical to those of a system with post-recording fringe rotation. In both cases, the required data are completely deterministic functions of time involving a small number of parameters.

Playback Equipment

(a) Delay Setting

The time of any sample, including the phase of the sampling clock, can be determined from information recorded in each header, along the lines described in VLBA Memo No 333. The playback system will supply samples at times specified by the correlator, or as close as possible in case the correlator wishes to track a point other than that observed. (These operations would be the same with post-recording fringe rotation.)

(b) Phase Switching

The playback system should complement the sign bit of each sample if the phase-reversed bit in the corresponding header is true. The header should contain sufficient redundancy so that this bit and other timing information in the header can be recovered in the presence of errors; but if an uncorrectable error is detected, the validity flag should be set false for the entire frame. It is expected that the playback system will have sufficient buffering so that the decoding of the header, setting of the validity flag, and complementing of the output sign will not require particularly fast computation.

Correlator Equipment

(a) Fringe Rotation

We assume that the fringe rate has been brought to zero at each VLBA station, so no fringe rotation equipment is needed. However, to accomodate foreign (non-VLBA) stations and to allow for special experiments, provision should be made in the station electronics for plugging in a "fringe rotation module" for each of the 16 channels. Normally, no module would be installed; but if installed, the module would multiply the samples from its channel by a 3-level quantization of sin(wt + p), where w and p are specified by computer commands. In this mode, one could either accept the SQRT(2) loss in SNR from DSB fringe rotation, or one could command the playback system to supply the same signal on two correlator channels, and command the corresponding two fringe rotation modules to maintain phase quadrature.

(b) Calibration Tone Extraction

Any calibration tones (pulses) inserted in the front ends will be frequency and phase shifted by the fringe rotation. In principle they could be extracted by undoing the fringe rotation, but the equipment would be unreasonably complicated. Therefore, we should consider dropping the pulse calibration system and relying instead on astronomical calibration, and on the stability of the electronics between calibrator observations. The variation of complex gain on any one channel should be dominated by the station clock (H maser), which would not be corrected by the pulse calibrator anyway; the differential gain between channels should be more stable than could be measured by the pulse calibrator.

(c) Cross-correlation

To obtain 512 frequency channels, only 1024 simple (real) cross correlators are needed per baseline, rather than the 1024 sine/cosine correlators (2048 simple ones) needed for the "traditional" VLBI correlator. The number of FFTs per second is also halved, and the computations required to combine pairs of FFT outputs are eliminated.

Consequences

(a) Phase Referencing

Here we restrict the term "phase referencing" to the case of measuring the visibility phases of two sources which are simultaneously contained in the single-antenna beams. Then, in principle, the same observation can be used to measure both. The trouble is that the sources are generally at significantly different delays and fringe rates (see Appendix B), so they cannot be simultaneously visible in the output of a single correlator. We discuss here the case of two sources, but everything is extendable in a obvious way to three or more sources.

Consider the following techniques: (1) At observe time, let half of the channels track one source and half the other. At correlate time, no adjustments are required but each source gets only half the bandwidth. (2) At observe time, alternate the tracking between the sources. (3) At observe time, track only one source on all channels, but use only 8 channels (at maximum channel bandwidth, this uses all the presently planned recording capacity). At correlate time, the playback system can supply each channel on two correlator input channels, with different delays appropriate to the two sources. The source which was not tracked at observe time will suffer the fractional-bit delay error, but it will be slowly varying and hence fully correctable. That source will also have a visibility phase which is varying at its residual fringe rate, so those correlator channels devoted to it will need to have a sufficiently high dump rate to follow these variations. (4) Process the tapes twice, once for each source. This is reasonable only if all 16 correlator input channels are needed for each source. The correlator would need a high dump rate on the pass for the untracked source. (5) Dump the correlator sufficiently fast to follow the untracked source, but also provide separate integrators for the tracked source (accumulating each dump) and the untracked source (accumulating each phase-corrected dump), along with sufficient computing power to do both simultaneously. This only works in the special case that the two sources have a small delay difference on all baselines ($BD < \sim K/8$ for bandwidth B, delay difference D, and K lags/channel; then $B =$ 64 MHz and $K = 1024$ give $D < 2$ μs, or 19 arcsec at 1 earth radius).

Methods (1) and (2) are wasteful of recording capacity, but are desirable in that they require no special correlator enhancements. Methods (4) and (5) require that the correlator have a fast dump capability on all channels, with an FFT and phase correction performed each dump time.

Method (3) looks like the best choice; the fast dump is required for only half the channels, and the full bandwidth can be utilized. Nevertheless, to accomodate possible future expansion of the recording bandwidth, we recommend that sufficient capacity be provided to process fast dumps from all channels.

For comparison, note that a system with post-recording fringe rotation would impose the same requirements on the correlator for methods (1), (2), (3), and (5). Only in method (4), with two tape passes, would the fast dump requirement be avoided. But if observations of this type are needed more than a small fraction of the time, then multi-pass processing is precluded.

We can quantify some of the above statements by using results from Appendix B. The worst case (worst sky position and latitude zero) derivatives of fringe rate and delay rate with respect to source position are 18.5 Hz/HPBW and .449 ns/s/arcmin, respectively, where HPBW is the half-power beamwidth of a 25 m diameter dish (uniformly illuminated). After fringe rotation for the beam center, the maximum residual fringe rate within the HPBW is 9.2 Hz. To track this, the correlator output should be dumped at least 4 times per cycle; thus a 40 Hz dump rate (25 ms) should be adequate, and less will suffice if we do not need to cover the full HPBW at all wavelengths. The delay rate difference shows that a source 10 arcmin off center would accumulate a delay error of 45 ns in a 10 s integration; this is less than 1 sample time (62.5 ns) at 16 MHz sampling rate.

(b) Foreign Stations

The most straightforward approach, and perhaps the least expensive, for incorporating non-VLBA stations is to insist that these stations provide fringe rotation with respect to the VLBA reference station, say the earth's center. This can be accomplished by their adopting VLBA baseband converters. It would be convenient for this and many other purposes if they would also adopt the VLBA station computer and some VLBA software. However, no specific hardware or software would be required, so long as the signal processing meets VLBA specifications. Phase switching would also need to be implemented, with each station being assigned its own Walsh function.

However, if some foreign stations were crucial to a particular experiment and were not equipped to VLBA specs, then it would be possible to use the optional station-based fringe rotation modules at the correlator. If this is done only for the foreign stations, with VLBA stations retaining fringe rotation in the LO, the resulting asymmetry of signal processing may produce some map artifacts; actually, we know of no reason to expect this, but such a configuration has not been tried. The option exists for using the post recording fringe rotators for all stations, but this will probably result in a severely reduced number of channels (or multi-pass processing), since we expect not to implement a full set of fringe rotation modules. Thus, we consider the correlator fringe rotators to be a backup or emergency-use-only system, with the primary method being to equip all stations to VLBA specs.

The special case of Mark III compatability deserves mention. We do not know whether it is feasible to modify Mark III electronics to include fringe rotation, variable phase sampling, and phase switching; if so, the cost of VLBA compatability for Mark III stations should be small. It seems unlikely that purchase of a complete set of VLBA electronics would be necessary. Note that compatability of the recording equipment alone is an independent matter which we are not addressing here.

(c) Bandpass Offsets

Unless some fringe rotation is applied prior to the baseband filters, the narrowest channel bandwidths now specified for the VLBA (namely 62.5 and 125 kHz) will be unusable at high frequencies, where the fringe rate approaches or exceeds the bandwidth. Even when the effect is less gross, a frequency offset which is a significant fraction of the channel bandwidth will destroy the bandpass matching among stations, leading to loss of "closure." Another way of saying this is that the complex gain on any baseline will be a function of fringe rate. In the extreme case of 86 GHz operation, where the fringe rate reaches 133 kHz (relative to earth center), the gain variation exceeds 1.6% even at our widest bandwidth (8 MHz). These problems are eliminated by doing the fringe rotation in an LO.

It has been suggested that only partial fringe rotation be done at the stations by offseting the LOs in discrete steps. With 10 kHz steps, the residual fringe rate on any baseline could be kept to less than 10 kHz. However, this would not eliminate the gain variation with fringe rate, and it would allow the residual fringe rate to pass through zero at many places in the (u,v) plane. Furthermore, it would require almost as much computation and bookkeeping by the station computer as would full fringe rotation. The gain variation at 86 GHz would still exceed 1% for bandwidths less than 1 MHz.

(d) Geodesy

We see no significant impact on geodesy experiments. The fundamental delay and phase observables would be available in the same way, regardless of where the fringe rotation is done. One small change is that the calibration of the phase difference between channels would depend on long-term measurements of astronomical calibrators, rather than short-term detection of injected tones; this should be very accurate, and is facilitated by having a dedicated array with stations that are nominally identical.

Conclusions

The main implication of this scheme is that the correlator is grossly simplified compared with the "conventional" correlator, in which fringe rotation is done for each baseline. We emphasize that the simplification is major, and not just a matter of saving a few cross multipliers. A great deal of the logic of the correlator may be devoted to implementing the fringe rotation, especially if it is done for each lag cell separately. The reduction in logic is then close to a factor of 3. In addition, the high speed computation of fringe phases and the distribution of this information are eliminated. Finally, calibration tone extraction is eliminated (although this occurs by default, and a fair comparison would eliminate it from the fringe-rotating correlator also). The net cost saving is estimated to be $495k (see Appendix A).

Some computational burden is shifted from the correlator to the monitor/control computers. Since the latter are distributed among the stations, the burden on any one computer is expected to be small, and well within its available capacity.

Thus, the advantages of the signal processing considered here seem to be: (1) the correlator is greatly simplified; (2) the baseband filters at different stations appear at the same source frequency, eliminating closure problems which would otherwise be severe at small bandwidths; (3) the small loss in SNR due to quantization in the fringe rotator is eliminated.

There appear to be these disadvantages: (1) it is somewhat more difficult to include foreign stations which do not have equivalent capabilities; (2) phase switching is required, and would cost something to implement [but phase switching is desirable even with post-recording fringe rotation, unless we are willing to discard data taken at low fringe rates]; (3) the fringe rotators add to the cost and complexity of the baseband converters, and to the load on the station computers; (4) changing the phase reference position requires fast correlator dump rates, even if separate tape passes are used [but covering more than one position in a single pass requires fast dumping even for post-recording fringe rotation, and this may be required].

Other consequences that have been mentioned as disadvantages but which we believe are easily overcome are: (1) there must be strict accountability for the fringe rotation which was done at observe time [we note that the same accountability is required if it is done at correlate time]; (2) high fringe rates after the final LO tend to reject the image of that LO (see Rogers, VLBA Memo No. 327) and this rejection is lost here [we note that the image rejection specification on the final mixer can be tightened to better than 30 dB].

ADDENDUM

After the above memo was complete, B. G. Clark pointed out that the requirements for updating the fringe rotation oscillators are not so stringent as was stated here. If only phase and rate are programmable, then use of the best linear fit to the computed phase over the interval between updates (rather than extrapolating the initial phase at the initial rate, as assumed above) leads to a worst case update interval of 0.14 sec (using earth center reference, 86 GHz, and peak phase error of 0.1 radian). This is within the capacity of the monitor/control bus (17 kbaud if 0.1 sec update interval is used, out of 56 kbaud capacity). However, to preserve a large margin in M/C bus capacity and to keep down the load on the station computer, we still recommend implementing oscillators whose phase acceleration is also programmable.

APPENDIX A: Cost Estimates

The latest available detailed budget for the correlator is given in Volume III of the VLBA Proposal (not formally published). Taking only the fabrication budgets, we find that eliminating the phase calibration extractors from the station electronics will save \$96.02k. In the correlator electronics, we delete the "phase processor" and reduce the "correlator" and "correlator accumulator" by a factor of 2, for a total saving of \$415.21k. This does not include the savings from elimination of the fringe rotation logic, which may be substantial. It also does not assume any reduction in the "correlator controller" or in either of the two fringe processors, even though our maximum proposed dump rate of 40 Hz is much less than the 160 Hz on which this budget is based. Nor have we included any reduction in development costs or software costs. Considering that the delay functions and the playback machines are now part of the recording system, the total correlator budget from Volume III is \$3006.8k, which becomes \$2495.6k after the simplifications considered here.

The electronics needed to implement fringe rotation in all 16 baseband converters at each station will add some cost. Assuming that the fractional-N synthesis technique is used (see Moffet, VLBA Memo 276), we estimate 10 digital MSI chips per converter will need to be added. This is 1600 chips for 16 converters and 10 stations; conservatively estimating \$10/chip, the total cost is \$16k. We are unable to estimate the cost of software to drive the fringe rotators, but we assume that it will be at least cancelled by the elimination of such software from the correlator.

The net estimated saving is thus \$495k; we think this can be taken as a lower bound.

APPENDIX B: Calculations

The geometrical delay on a baseline $\vec{B}$ for source position $\vec{s}$ is

(1) $D = \vec{s} \bullet \vec{B}/c,$

where, in convenient coordinates,

(2) $\vec{B} = r_E \, (\cos L, \, 0, \, \sin L)$

$\vec{s} = (\sin h \cos \delta, \, \cos h \cos \delta, \, \sin \delta).$

Here we have taken one end of the baseline to be at the center of the earth, with r_E the earth's radius and L the station's latitude. Then

(3) $D = r_E \, (\cos L \cos \delta \sin \omega_E t + \sin L \sin \delta)/c$

where $h = \omega_E t$. The fringe phase and its first two time derivatives are

(4) $\phi = 2\pi \, D \, f = 2\pi \, D \, c/\lambda$

$= (2\pi \, r_E/\lambda) \, (\cos L \cos \delta \sin \omega_E t + \sin L \sin \delta)$

(5) $\dot{\phi} = (2\pi \, r_E \, \omega_E/\lambda) \cos L \cos \delta \cos \omega_E t$

(6) $\ddot{\phi} = -(2\pi \, r_E \, \omega_E^2/\lambda) \cos L \cos \delta \sin \omega_E t$

Differentiating $\dot{\phi}$ and $\dot{D}$ with respect to hour angle and declination gives:

(7) $\frac{\partial \dot{\phi}}{\partial h} = - \, (2\pi \, r_E \, \omega_E/\lambda) \cos L \cos \delta \sin \omega_E t$

8) $\frac{\partial \dot{\phi}}{\partial \delta} = -(2\pi \, r_E \, \omega_E/\lambda) \cos L \sin \delta \cos \omega_E t$

(9) $\frac{\partial \dot{D}}{2h} = -(r_E \, \omega_E/c) \cos L \cos \delta \sin \omega_E t$

(10) $\frac{\partial \dot{D}}{\partial \delta} = -(r_E \, \omega_E/c) \cos L \sin \delta \cos \omega_E t$

To evaluate the worst case, let all trigonometric factors be unity; then, using $r_E = 6378$ km, we find that the fringe rate derivative is (464 m Hz/rad)/λ, and the delay rate derivative is 1.544 μs/s/rad. Results quoted in the text follow directly from the latter numbers.