Submillimeter Array Technical Memorandum

Number 83
Date 3/24/95
From .... Andrew Dowd

Implementation of The Secondary Phase Rotator

I. Introduction

The present architecture of the SMA downconversion chain requires a second phase rotation. This processing step is performed in the final LO to remove residual phase errors due to geometric delay tracking. The effect of this unit is a slowly varying phase shift. Physically, this phase shift is modulated on the LO, but it is applied to the astronomical data signal during the final downconversion stage. Appendix A gives some background on the requirements for this processing step.

The body of this document will present the proposed implementation of the phase rotator. This including details of the proposed design. Another section will discuss a requisite calibration procedure. To achieve acceptable performance, the proposed phase shifter requires this calibration. As it turns out, the calibration procedure became a more formidable problem than the actual phase modulator.

II. Specifications

Some of the quoted specifications relate to operational requirements, while others are tied to performance issues.

Secondary Phase Shift Specifications

	Specification	Actual (So Far)
Phase Rate (υ) (Max)	1.47 Hz (min)	7.9 Hz
RMS Phase Error (Initial)	1" (max)	($\upsilon < 0.12$ Hz) 0.05°, ($\upsilon > 0.12$ Hz) Eqn 2
RMS Phase Error (After 1 Year)	1" (max)	($\upsilon < 0.12$ Hz) 0.08°, ($\upsilon > 0.12$ Hz) Eqn 2
Amplitude modulation (LO)	-	0.01 dB
Calibration Frequency	Once per Year	-
Calibration Time per Freq.		10 minutes
EEPROM Bytes per Freq.		128
Units Needed for SMA	288 (See Note 1)	Prototype Built

Notes

1. In practice, the SMA could be built with less units by sharing LO signals between chunks. However, any cost savings would be largely diluted by cabling costs.

III. **Implementation**

This device applies a variable phase correction to the final LO in the SMA. In general this LO can take on 4 different values in the frequency band of 700 MHz to 1.3 GHz. Four frequency settings are used to downconvert different chunks. However, it is conceivable that different frequencies settings will be used to implement unusual observing modes. It will be assumed that the operating frequency for a given LO will consist of a small set of fixed frequencies that will be changed infrequently.

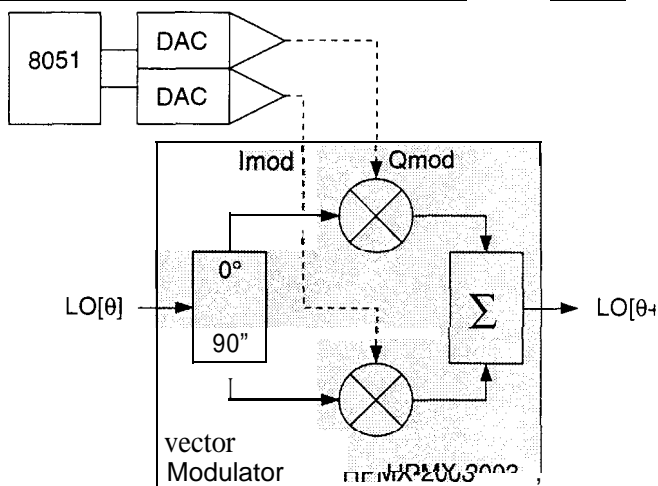The phase modulation is directly applied to the LO signal using a *Vector Modulator* as outlined in Figure 1. The modulator accepts two control signals (I and Q) which drive two mixers in quadrature. If the I and Q inputs are simply DC voltages the transfer function is shown in Figure #2 (which was measured using an actual prototype). If the vector modulator is ideal, then applying a pair of quadrature CW signals to the modulation ports would produce a frequency offset (phase rate) equal to the modulation frequency. In Figure #2, this can be visualized as swinging around a circle (or contour) on the power curve and ascending a continuous spiral on the phase plot. If this is done perfectly, power gain will be constant and phase will increase (rotate) smoothly.

In practice, there are several problems. First, although it's not obvious from the measured plots, there are small distortions in the transfer function of the device. Even with perfect modulation signals, the resulting LO would exhibit phase errors and amplitude modulation. (These errors are equivalent to images produced by a single-sideband mixers). The errors produced for a good vector modulator are around 0.25 dB AM and approximately 1 degree of RMS phase error.

In and of itself, these errors are marginally acceptable. There is no hard specification on
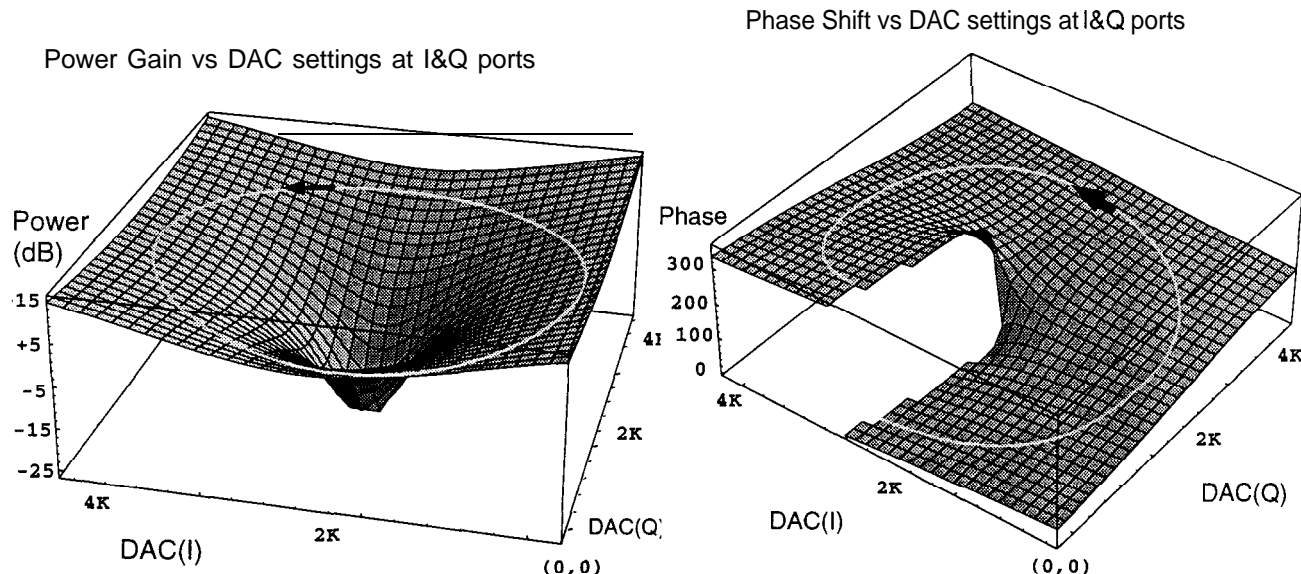

Figure 1 – Phase Rotator

Figure 2 – Measured plots of power gain and phase versus DAC settings.

maximum AM for the astronomical signal. However, any amplitude modulation introduced by the LO will NOT be directly reflected in the signal path. If the mixer is properly saturated, amplitude variations will be attenuated. On the other hand, LO phase errors will be directly reproduced as errors in the signal phase.

Even if the quoted numbers of phase imbalance were acceptable, there is still the problem of generating a "perfect" modulation signal (in quadrature and 288 times!). Also, the quoted numbers are for a the vector modulator operating over a relatively narrow bandwidth. As a consequence, our LO range could not be covered by a single device. Fortunately, there's a solution that fixes these issues, and also produce a vast improvement in the errors possible using a practical vector modulator.

Before describing the proposed solution, it's useful to consider a related method of generating the modulating signals using a DDS (Direct Digital Synthesis). The signals applied to the I and Q ports of the vector modulator must be quadrature sine waves with very fine frequency resolution (≈ 1 mHz) to produce the different phase rates. This type of job is normally done with a moderately advanced DDS. In practice, a DDS would be expensive, and increase the errors in the system due to it's own imperfections. However, by constructing a discrete DDS circuit using DACs (Digital-to-Analog converter) and a micro, its possible to measure and correct any errors in both the generating circuit and the vector modulator. Thus, instead of trying to generate a very good CW signal and hoping that it will satisfy an imperfect vector modulator; the proposed solution will a create quadrature signal with offsets which tend to counteract errors in the vector modulator. The proposed solution is, in effect, a very slow DDS with adjustable Sin tables.

The DAC based dual modulation generator (sketched in Figure 1) will generate a quantized stair-step function to approximate the phase gradient. The Microprocessor has a table of DAC values which were derived from a calibration procedure (more on this later). At fixed time intervals (locked to the array clock), the DAC will be programmed to it's next step position (See Figure 3). Even with perfect calibration of the circuit, there will be a phase error imposed by the DAC resolution. The resolution limits apply to both the stepsize of the DAC and update rate. The phase update rate has been set to 1msec (which is limited by the microprocessor load and not the DAC settling time). The DAC has 12 bits, which translates into approximately 0.04" per step. Therefore, below a

RMS Phase Error

1. Phase Rate(v) $> 0.12 Hz$

$$\Phi_{total} \approx \sqrt{\left(\frac{\upsilon}{7.9\,Hz}\right)^2 + \left(\Phi_{cal}\right)^2} \qquad (1)$$

2. Phase Rate(v) $< 0.12\ Hz$

$$\Phi_{total} \approx \sqrt{\left(0.015°\right)^2 + \left(\Phi_{cal}\right)^2} \qquad (2)$$
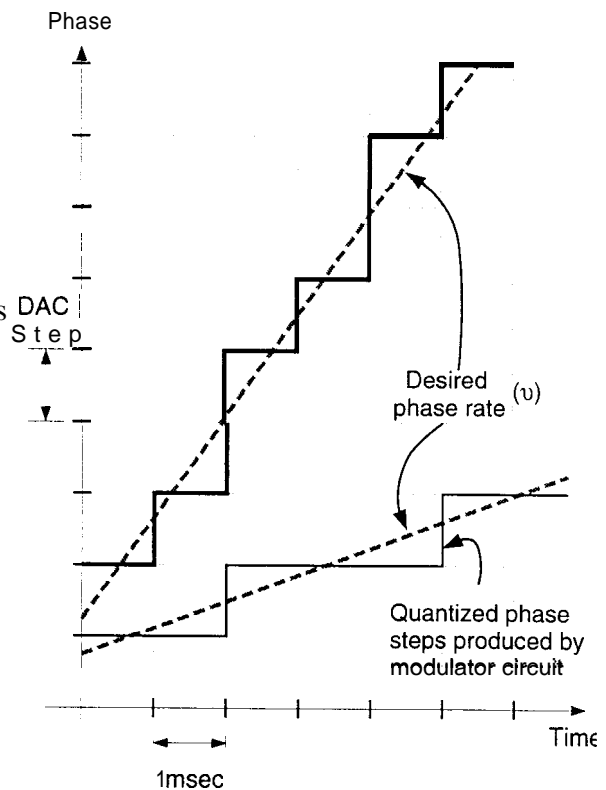

Figure 3 - Arbitrary phase rate generator

phase rate of 0.12 Hz, the DAC never increments by more than one step. In this case, RMS phase error is limited by the DAC resolution plus any instrumental error which remains after calibration. On the other hand, when the phase rate exceeds 0.12 Hz, the DAC will be forced to jump several increments in a timing cycle. This will degrade phase noise performance. Phase noise will increase proportional to the phase rate. (An estimate of phase error is given in equations 1&2).

Amplitude modulation will not increase with phase rate because the power should be the same at each DAC setting. Therefore AM should will determined by the instrumental amplitude error.

The required maximum phase rate for the SMA is 1.47 Hz, which will impose a phase error of 0.186" RMS (excluding calibration error). Significantly faster phase rates would require a faster update rate.

One potential advantage of the DAC modulator is it's ability to generate non-linear phase rates. The only limitation is the processing capabilities of the 805 1 micro to determine the DAC values. However, this capability will probably not be necessary for the SMA correlator because of it's rapid dump rate. This makes linear phase rates quite accurate.

In practice, the performance of this system could be severely limited by the accuracy of the calibration and it's long term stability. This issues will be discussed in the following section. By calibrating the device, it's possible to utilized a single vector modulator to cover all the expected LO frequencies. In a conventional modulator design, this would be difficult because of bandwidth limitations in the 90 hybrid (See Figure #1). The proposed vector modulator is the HPMX-2003, which is only characterized for operational from 800- 1000 MHz. However, the calibration allows operation over the full bandwidth (approx 750 MHz - 1.25 GHz). There is some modest loss in gain and a slight increase in parameter sensitivity outside the devices typical operating band, but these effects are small and manageable.

IV. Calibration

The task of calibrating this circuit involves finding the DAC settings which will generate points on the equal power contour of the devices transfer function (See Figure 2). In this process the DAC circuit and the vector modulator become paired devices because the calibration absorbs errors from both. The goal is to produce a table (or model) which can be stored in EEPROM. The nonvolatile storage is necessary because the proposed calibration scheme will be performed off-line. This raises issues of calibration stability which must be tested. Also, given the number of devices involved, the required calibration time is quite important.

The brute force approach to calibrating this device would be to create a table of DAC settings which is optimized for a given phase step and at equal power gain. To achieve low phase error, the number of entries in the table must be fairly large (1 degree steps = 360 entries). Each entry value must be individually optimized over a two-dimensional field of the I and Q DAC values. The phase and power is measured with a network analyzer.

The implementation of this calibration scheme is sketched in Figure 5.   In the first experimental attempts at calibration, the brute force method was automated and produced good results. Unfortunately, even with complete automation it took a rather long time to calibrate a given device for a given frequency. In fact, it took 4 hours, which means the calibration of all the SMA phase rotators would take a total of 576 working days (288 rotators x 4 freq values x 4 hours / 8 hours per Day). Undoubtedly, the algorithm to perform the optimization could be speed up, but it seemed clear to me that some major shift in methodology was needed. So I changed direction at this point and starting

working on a model which could be used to predict the vector modulator response.

To the first order, a reasonable model of vector modulator response could be used to accelerate the brute force search algorithm. However, if the model was reasonable good at predicting response, it could also be used to parameterize the calibration.   Then the calibration procedure would optimize the model parameters. A model based approach could greatly reduce the size of the required nonvolatile storage. Also, it could improve phase noise because the quantization of the data is no longer limited by the size of the data table.

As mentioned earlier, the goal of our model is an estimate of the I and Q DAC values that lie on the equal power contour for a given phase.   In the ideal case, the DAC model can be described with the following equations:

$$DAC_I = A \; Sin(\phi)$$

$$DAC_Q = A \; Cos(\phi)$$

(3)

Where A = a scalar term, ϕ = the desired phase

A number of practical issues will induce deviations from this model. For example, there will be some phase error in the 90 hybrid. Also, the gain between the Inphase and Quadrature legs will be different. Finally, the will be some DC offset which could differ between the I and Q ports. In a symmetric way, similar errors will exist in the DAC circuit. And of course all these deviations will be frequency dependent.   By incorporating the known error terms produces a more complex model of this circuit:

$$DAC_I = A_I \; Sin(\phi + \varepsilon_I) - X_I$$

$$DAC_Q = A_Q \; Cos(\phi + \varepsilon_Q) - X_Q$$

(4)

Where X = DC offset term, A = scalar gain term, ε = phase offset term

The new model of complex gain would require much less time to calibrate than a full table oriented search. With only 3 unknowns (per axis), this model could be fit with only 3 measured data points. Unfortunately, the residual error using this model were empirically determined to be too large. From the error residuals, it was clear the problem was systematic in nature. Therefore, a modified model was proposed where equation 4 was fitted by quadrants. Thus, instead of a single value for A,ε, and X (per axis), the new model has 4 sets of parameters. To determine the model parameters, eight data points are found as indicated in Figure #4. These 8 experiment data points are located by moving along a radial line until the power gain is correct. At each measured point, the power gain, phase, and DAC settings are saved to disk. (The measured power and phase is used in the parameter estimation to improve the fit.) The fitting was performed with equation 5. With three data points per quadrant, the parameter fit is deterministic. The measured phase values are inserted in ϕ. Due to quantization, optimum "DAC" values are never
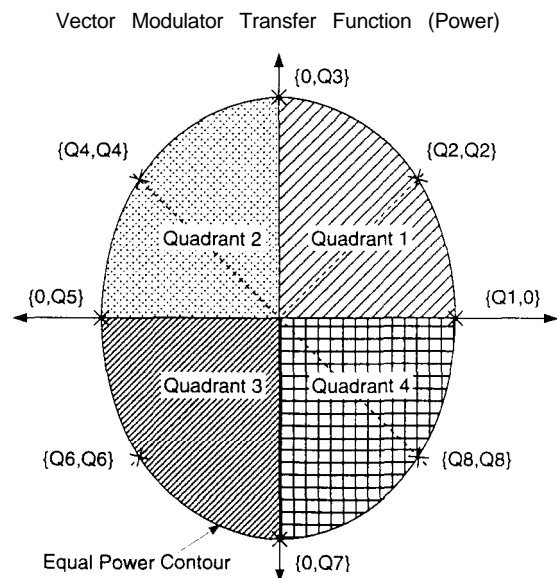

Figure 4 – Model Quadrants and Measured Points

$$\begin{bmatrix} cos\,\phi_1 & sin\,\phi_1 & 1 \\ cos\,\phi_2 & sin\,\phi_2 & 1 \\ cos\,\phi_3 & sin\,\phi_3 & 1 \end{bmatrix} \begin{bmatrix} A\,cos\,\varepsilon \\ -A\,sin\,\varepsilon \\ -X \end{bmatrix} = \begin{bmatrix} DAC_1 \\ DAC_2 \\ DAC_3 \end{bmatrix} \tag{5}$$

quite on the equal power contour, so to improve the fit, the DAC values are interpolated from the measured data points.   This complexity is probably unnecessary, but should improve phase noise slightly with no time penalty in the calibration cycle (the calibration time is dominated by measurement time and not calculation). The equation is then solved for {A,X and ε}, which become the model parameters.

This operation is performed for each quadrant on both axis. (It's worth mentioning that this model does not include any cross-terms between the I and Q axis.)   The end result is a table of 24 parameters to describe a given equal power contour. In practice, one additional term is needed to indicate the boundaries of the quadrants. Consider the following question: If you want to produce a given phase value, which quadrant model do your use? This could have been constrained in the measurement process, but it was easier to simply allow the quadrants to be defined as a parameter. Thus the total model contains 32 data values (128 bytes). This is quite tractable. A modest sized EEPROM of 5 12 bytes could handle all the parameters for a given modulator and the four primary operating frequencies of the array.

From experiments, this model gave more than acceptable performance. Figure 6 gives a plot of error residuals after calibrating a phase rotator using the quadrant based model. These plots were taken by commanding the device to a specific phase values, then measuring the resulting amplitude
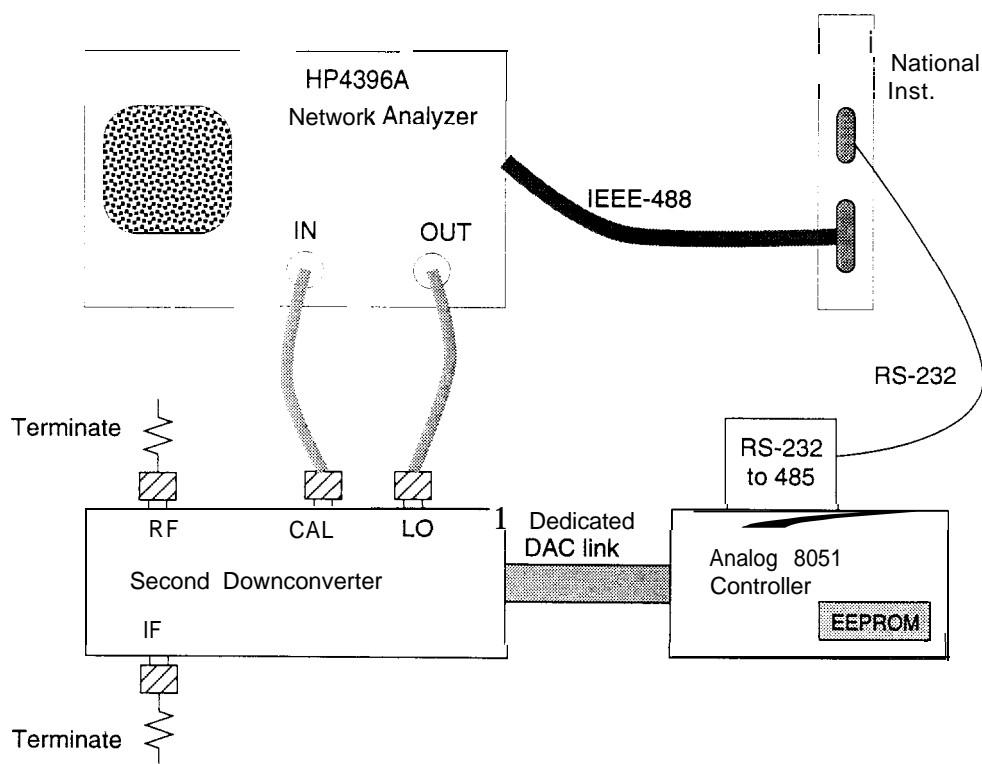

Figure 5 - Calibration setup

and phase. Any offset is removed. The worse case phase deviation was well below ±0. 1 degrees and amplitude modulation below ±0.03 dB.

In another attempt to improve calibration time, I also implemented a more intelligent algorithm for guessing power along a radius. The algorithm measures two points, one at maximum power and the other halfway to the center. Although it's not obvious from Figure 2, the power transfer function is roughly linear in voltage, so converting the measurements to voltage, its possible to do a linear interpolation on the two values to give a good guess as to the correct DAC setting for a given gain. By iterating this method, the equal power contour will normally require only 4 measurements to locate the optimum DAC setting.   Between this refinement and the model based approach, the total calibration time was reduce to a more manageable 10 minutes.

One final matter is the long term stability of the calibration. The calibration procedure as sketched in Figure 5 must be performed off-line using a network analyzer and a VXI based-computer. Re-calibration of a module will require taking it out-of-service.   From the point of view of operations, it's important that re-calibration be done as infrequently as possible. Thus, the long-term stability of the calibration must be quite good.   At the very minimum, I'd like this calibration to be performed no
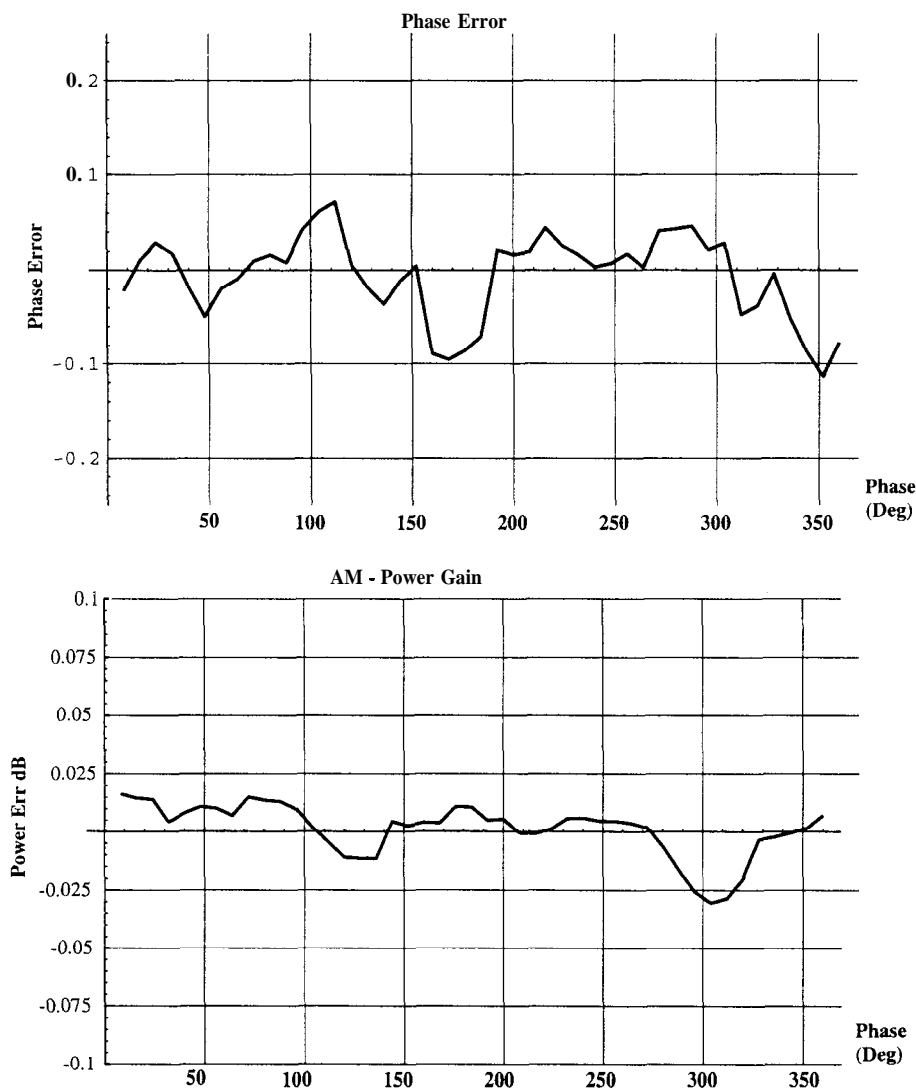
Figure 6 - Instrumental phase and power(gain) errors immediately after calibration

more than once per year. To this end, I've begun a long-term test on a prototype.

At this time, the test module has been running for about 4 months. During most of this period the power was applied, but it was also off for about 1 month. The error residuals are plotted in Figure 7. Not surprisingly the error has increased. No attempt was made to control the temperature, so some of this residual could be due to drifts. The measured errors are tolerable, but could be even smaller when the unit is operated in the temperature-controlled correlator room. This also means some reasonble effort should be made to perform the calibration at the same temperature as the correlator room.

V. Conclusion

The proposed secondary phase rotator will use 288 independent vector modulator at the last downconversion stage. Experiments on the proposed design indicate acceptable performance if an off-line calibration procedure is applied to the circuit. The measured calibration parameters are placed in EEPROM for long-term storage. The actual calibration will require only a few minutes to perform.
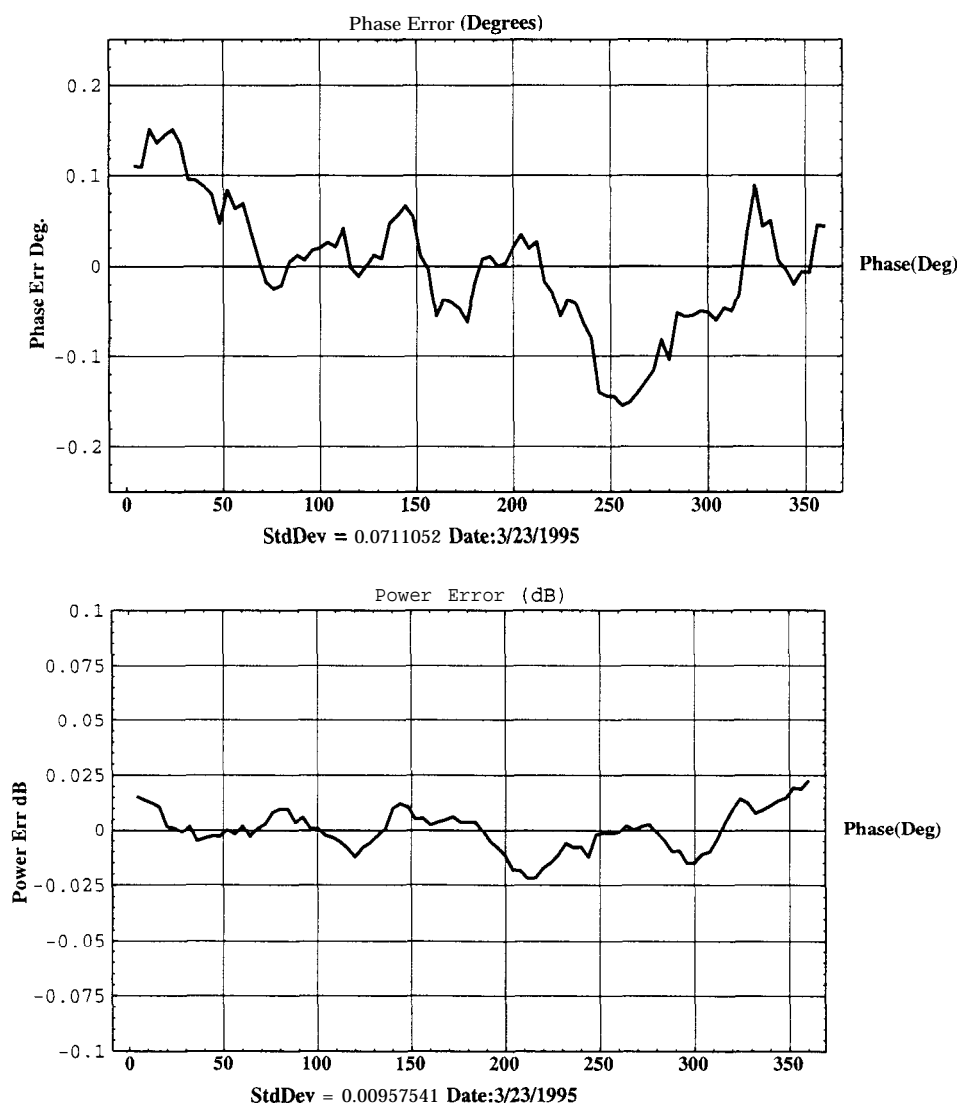

Figure 7 – Phase and power (gain) errors after 4 months (applying original calibration parameters)

A final question is the long-term stability of the calibration. Initial indications look good, but an exhaustive study on this is impractical. Validation on this matter will require some time and experience.

VI. **References**

[1] *Interferometry and Synthesis in Radio Astronomy,* Thompson, et al., 1991

[2] SMA Technical Memo #68, Colin Masson

[3] SMA Technical Memo #56, Colin Masson

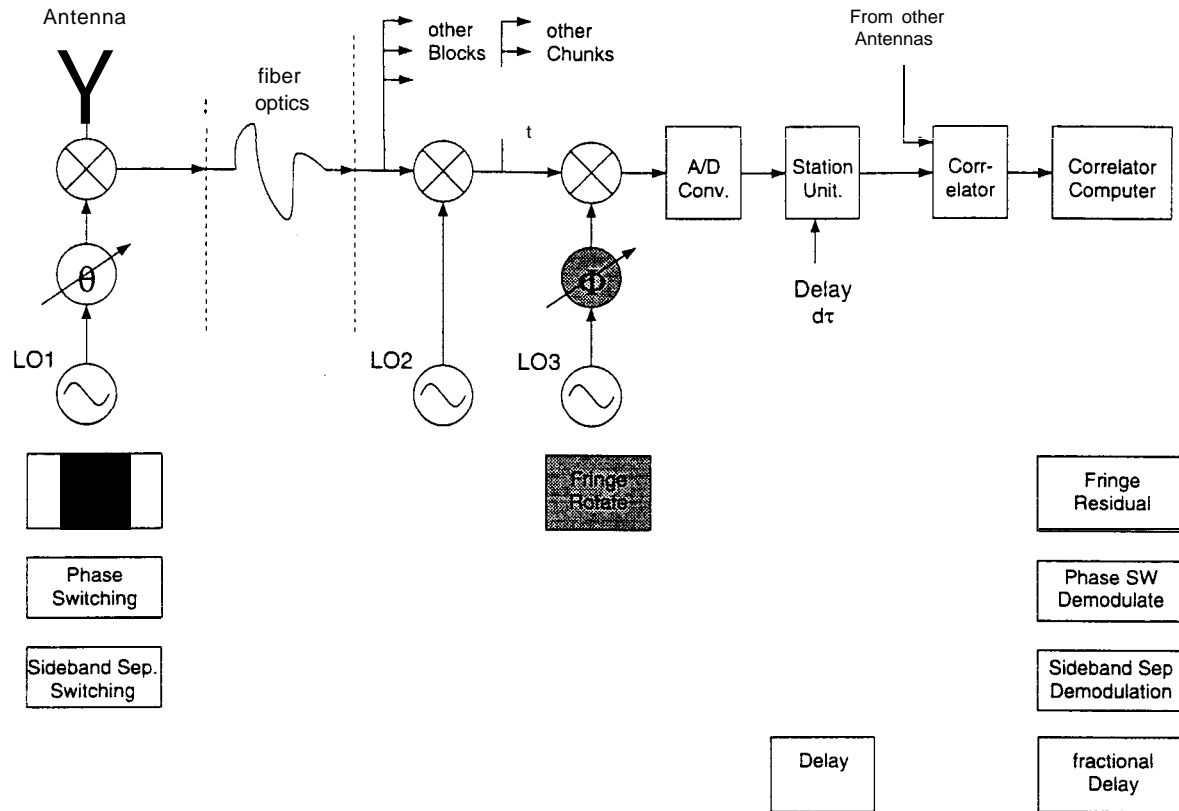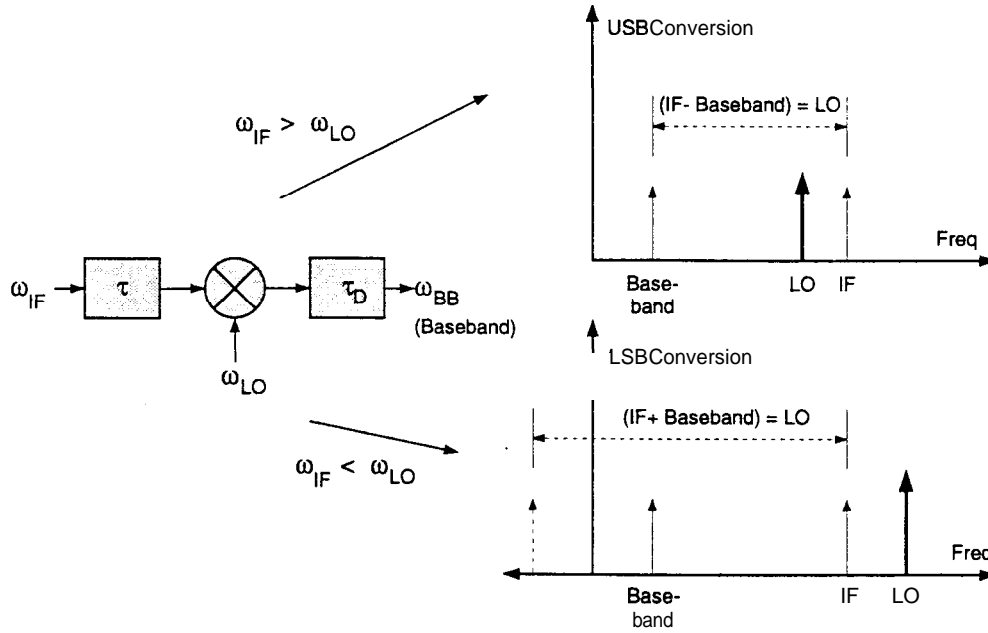[4] *Etude De La Carte,* Internal IRAM Technical Memo, Marc Torres

Figure Al - Signal Processing steps in the SMA

Appendix A. **Secondary Fringe Rotation, Necessary Specifications**

        Many of the specifications listed in Table #1 are related to operational issues. However, two of the quoted values are tied to the performance of the SMA. They are the maximum phase rate and the RMS phase error. Necessary background on issues can be found in SMA technical memos 68 and 56 by Colin Masson. This appendix will assume knowledge of these documents. For reference purposes, figure Al gives a revised sketch of the signal processing steps in the SMA. The secondary phase (or fringe) rotator is highlighted in grey.

        The specification of RMS phase error is a relatively straightforward matter because any error at this stage will have a simple one-to-one equivalence to the data which enters the correlator. Thus, a 5 degree phase offset in the secondary phase rotator will cause a 5 degree phase offset in the measured data. Any phase noise (or error) introduced in this stage must be small enough to avoid excessive losses in correlation. Since there are other phase deviations in the system, the allowable phase error will be somewhat less than for the total allowance for the full array. An RMS phase error of 1 degree meets these goals, and would introduce of loss of only 0.03%

        A slightly more complicated issue is the maximum phase rate of this unit. This requires delving into the purpose of this correction to the array's signal processing. Broadly speaking, the purpose of the phase rotator is to compensate for the residual phase shifts which remain after applying the geometric delay compensation in the digital data stream. By trying to remove a physical delay of the astronomical signal in the baseband signal, there remains a residual phase shift, because the two delays are being applied at different frequencies. As the geometric delay changes, the magnitude of the residual phase changes, thus it is necessary to track the phase correction. Our goal for this discussion

$\omega_{IF} > \omega_{LO}$

USBConversion

(IF- Baseband) = LO

Freq

Base-band   LO   IF

$\omega_{IF} \rightarrow \boxed{\tau} \rightarrow \bigotimes \rightarrow \boxed{\tau_D} \rightarrow \omega_{BB}$ (Baseband)

ω_{LO}

LSBConversion

$\omega_{IF} < \omega_{LO}$

(IF+ Baseband) = LO

Freq

Base-band   IF   LO

Baseband Signal (USB conversion) =

$$a_U \cos[-\omega_{IF}\tau - (\omega_{IF} - \omega_{LO})\tau_D + (\omega_{IF} - \omega_{LO})t]$$

$$a_U \cos[-\omega_{IF}\tau - (\omega_{IF} - \omega_{LO})(-\tau) + (\omega_{IF} - \omega_{LO})t]$$

$$a_U \cos[(-\omega_{LO})\tau + (\omega_{IF} - \omega_{LO})t]$$

Baseband Signal (LSB conversion) =

$$a_L \cos[-\omega_{IF}\tau - (\omega_{IF} + \omega_{LO})\tau_D + (\omega_{IF} - \omega_{LO})t]$$

$$a_L \cos[-\omega_{IF}\tau - (\omega_{IF} + \omega_{LO})(-\tau) + (\omega_{IF} - \omega_{LO})t]$$

$$a_L \cos[(\omega_{LO})\tau + (\omega_{IF} - \omega_{LO})t]$$

Figure A2 - Single conversion example: USB and LSB

is determining how fast this phase rate will change in the worst case operation of the SMA. (In general the phase rate resolution would also be an issue. However, the proposed phase rotator has no significant limit on phase rate resolution.)

Ultimately, the required phase rate is tied to two things. Since phase equals delay times frequency, one term is the geometric delay. The second term is related to the frequency translation. The rate of change of the geometric delay is a well known value quoted in SMA memo #68: Max delay rate = 242.4 ps/s. This rate is for a 1 km baseline.

The term related to frequency translation is not so straightforward. From Figure A 1, we see that two stages of downconversion are being applied between the receiver and correlator (The receiver LO has an independent phase rotator). Thus, the secondary phase rotation must compensate for residual phase from both stages of conversion. Not only must we consider the cumulative effect of two stages, each stage can be either an upper sideband (USB) or lower sideband (LSB) conversion. In fact all four possible combinations of conversion types exist in the present implementation.

To warm up to this calculation consider the problem of determining the residual phase in a single stage downconversion from IF to Baseband. This translation can be done with upper or lower sideband mixing. If we fix the IF and Baseband frequencies to be the same in both conversions, we get the situation in figure A2. In this arrangement, the LO's will differ, with the LSB LO above the IF frequency and the USB LO below the IF frequency. We'll assume the geometric delay is perfectly corrected with an equal and opposite digital delay (OK negative delay is a bit of a stretch, but I never promised a casual system). The resulting residual phase which must be compensated for with the

phase rotators is set by the delay size and the frequency of the *LO*. For LSB conversion the phase will be larger because it's LO is higher. Thus, although these conversion do effectively identical operations in term of frequency translations, the phase rotation (and rate) is different!

I stated that the residual phase is proportional to the LO frequency in the hypothetical situation. Well, looking at it another way, the residual phase is proportional to the frequency difference from RF to IF. Ah, but wait a minute, it's easy to see this for the USB case where $F_{LO} = F_{RF} - F_{IF}$, but that's not true for the LSB. Well, in fact it is, if we consider that LSB conversion causes the IF to get flipped over to negative frequencies. (first negative delay, now negative frequencies: What next, tachyon emission!) . Therefore, the frequency difference for the LSB conversion is $F_{LO} = F_{RF} - (-F_{IF})$.

Now to untangle the multiconversion case, we need to worry about IF's that end up at "negative" frequencies, because they increase the residual phase and thereby increase the phase rate. I won't go into the gory details of why, but it is possible to ignore the particulars of the downconversion. All we really need to consider is the initial IF frequency, it's eventual location at baseband and whether it resides at negative frequencies. The last question is a little confusing, but there's simple test. It will occur when chunk are "reversed". When a lower sideband conversion throws things into negative frequencies, it reverses the frequencies. Thus any chunk which has it's frequency order reversed relative it's IF can be thought if as a "negative" frequency chunk. Thus, phase residual will equal the IF frequencies plus the baseband frequency. In the other case, the phase residual is the IF frequency less the baseband frequency.

All of these details will be essential to the array software, which must generate the necessary phase rate. But at the moment, all that matters is the worse case. The worse case is when the frequency separation is a maximum. In the proposed downconversion scheme this occurs at IF frequency 5.984 Mhz and Baseband frequency of 112 MHz. This case is for a reversed baseband, so the maximum phase rate is given as follows:

$$v_{max} = \mathbf{242.2} \, \frac{psec}{sec} \, \mathbf{(5.984} \, + \mathbf{0.112} \, GHz)$$

$$= 1.48 \, Hz$$