ALMA Calibration

Bryan Butler, Stephane Guilloteau, Al Wootten, Ewine van Dishoeck

2003-Aug-07

1 Introduction

This document outlines the various quantities that will need to be measured or taken into account in order to collect and calibrate ALMA data. It is an update of Chapter 3 of the old Project Book, and supersedes it. Traditionally, "calibration" has often been thought of as strictly a post-processing exercise in radio interferometry, essentially only involving things done to data after it has been collected. In this document we take a broader view, and include all quantities that must be measured *before* correlation of the antenna signals as well. These are still formally "calibrations", since they are measurements of instrumental parameters – they are generally just measured less frequently. They are, however, no less important than the post-processing calibrations. In addition, we address some topics that are not even direct measurements, but are rather things that affect the measurement of our desired quantities. An example is the relativistic deflection of radio waves in the gravitational potential of the sun, which is not really a directly measured or calibrated quantity (except indirectly), but does certainly affect our ability to properly calculate delay, which in turn affects our ability to calibrate, for example, antenna station locations.

Accounting for all of these types of calibrations, measurements and effects is critical for ALMA to achieve its full potential. We must understand what effects must be accounted for, and how we will measure and/or correct for them during the data collection and post-processing. Never before has a radio astronomical instrument been built with such a detailed understanding of the site and its impact on the telescope. With this knowledge in hand, we can optimize the full measurement and calibration strategy to produce the maximum scientific output for ALMA.

In addition to simply describing the different types of calibrations, measurements and effects anticipated for ALMA, we also provide some specifications on the accuracy to which the measured quantities must be determined or effects must be accounted for. As for other interferometric arrays, the necessary or possible calibrations can be split into several types (see, e.g., Fomalont & Perley 1999; Thompson, Moran, & Swenson 2001). For our purposes we divide the calibrations into two main types:

- Precalibrations. These are quantities that must be known when the data is collected, and are used to make corrections to the signals or data at that time. They are by definition irreversible. They generally vary relatively slowly (longer than a single observation, and typically days to weeks. They are often fundamental instrumental quantities, collected during special (and relatively infrequent) calibration observations. In some instances, however, they vary more rapidly, and may have to be measured during the observation.

- Postcalibrations. These are quantities that can be measured before, during, or after the actual observation, which can be applied to the visibilities or total power data at any point in the data postprocessing. This is the majority of what is normally considered "calibration" historically.

Note that a special case of postcalibration is selfcalibration. If the target source is strong enough, then the technique of selfcalibration (see, e.g., Cornwell & Fomalont 1999) can be used to derive some calibration quantities. We do not treat the topic of selfcalibration in this document, because we cannot assume that in all cases it can be used for ALMA observations.

The calibration requirements in this document all derive from a series of meetings, discussions, and memos on the requirements necessary for the types of scientific experiments anticipated to be done with ALMA, and incorporate input from the ALMA Science IPT, other ALMA IPTs, the ASAC, and the community at large. They also incorporate a sense of reality, when the desired accuracy is simply not achievable or is too expensive. These scientific experiments are described briefly in the Requirements section, and examples given where appropriate.

In some cases, achieving the required accuracy requires the use of special measurement devices, either on the antenna, or separate from it. These devices are described where necessary, and requirements on their ability to measure the necessary quantity are given.

Finally, some of these calibration techniques require development and testing, as they are not commonly implemented on existing arrays. In these cases, there will be an attempt to test the technique on an existing array or antenna (30-m, OVRO, BIMA, PdBI, VLA, etc…), or on the prototype antennas and interferometer at the ATF. In some cases, testing on the early ALMA array itself will be required.

2 Requirements

The calibration requirements for ALMA derive from the types of experiments foreseen, and their required accuracies. In general, the astronomers want things "as good as possible", and we should always strive to make this so, but often it is simply not feasible to make things as accurate as desired by the astronomers. In some cases a direct cost increase is implied by improving a particular characteristic of the array, and that cost may be deemed as prohibitive. In others, it simply may not be technically feasible to reach the desired levels. We note here also that some desired final accuracy parameters are very hard to turn into specific calibration requirements, since ALMA will be such a complex instrument. In many cases simulations of the behavior can help, but in others we merely make educated guesses and try to make things as well as we can. This is not completely satisfying, but is the situation we are in. The various different requirements on calibration accuracy will now be described.

2.1 Antenna Characteristics

2.1.1 Pointing

Accurate pointing, or at least knowing *where* the antenna *was* pointed as a function of time, is important for most mosaiced observations. It is anticipated that a very large fraction of ALMA observations will be mosaiced – as much as 50% overall or even more in the compact configurations. It is also important for single-field submillimeter observations, given the very small primary beam at those frequencies. High frequency mosaics would benefit from better pointing than it is deemed possible to obtain with a 12m diameter aperture. Note that single-field observations at the lower frequencies, mosaiced observations at ν<120 GHz and mosaics which do not require high fidelity ("fidelity", though its precise meaning can be tricky, is essentially the ability to faithfully reproduce the actual sky brightness at all places in the image) at even higher frequencies will not require the best pointing.

In an ideal world, we would know the types of experiments to be performed exactly, along with specific accuracy requirements, and would specify a pointing requirement that allowed those experiments to be done (derived via simulation). Unfortunately, this is not the case, since the it is difficult for the astronomers to come to a consensus on the types of experiments which are most important to ALMA, and the derived requirement depends strongly on the types of experiments. In addition, it is difficult to define a single imaging metric which is applicable to all types of astronomical sources (see, e.g., the wide variety of imaging metrics reported in Pety et al. 2001b). We are not completely lost, however, as we do have some guidance for what the errors might be for at least some sources, and we have at least a general idea of the image quality that we want out (the disagreement is in the details). In general, it is agreed that much science demands an image fidelity (by some definition) of around 100 (i.e., errors in the image are 1%), and that this must be true for large, complex sources.

In addressing the general problem of required pointing accuracy for high fidelity mosaicing, Cornwell, Holdaway, & Uson (1993) derived a pointing requirement of 1/16th of the primary beam FWHM, which translates to 0.85 arcsec for a 12m diameter primary aperture at 0.8 mm wavelength. This was for a 40 element array (the old MMA), and to obtain an image fidelity (as they define it) of 100 on a particular simulated image. Holdaway (1997a) performed a more detailed analysis, showing the effects of pointing errors ranging from totally random to totally systematic. He found that with 12m antennas, a 1.1 arcsecond pointing error resulted in a median image fidelity of < 20 at 0.8 mm. Morita (2001) performed a similar study, and found that with 0.6 arcsecond pointing errors on 12m antennas, the image fidelity in a 7x7 mosaic was limited to <30 for ν>200 GHz. Adding an array of smaller antennas (the ACA) improved this by about a factor of two. Similar, but much more detailed, studies were performed in Pety et al. (2001a, 2001b), who obtained qualitatively similar results. All of these studies indicate that the old adage of "as good as you can make it" is certainly true – to quote Sebastian von Hoerner (in VLA Test Memo 129, 1981): "In summary, the astronomical demands are not so well-suited for deriving a well defined specification for the pointing error … Thus, it comes back to wanting the pointing error just as small as possible, and then to live with what one can get." This is as true for ALMA as it was for the VLA. In discussions with the prototype antenna vendors, it became clear that 0.6 arcsec was an acceptable error to them, and that is what has been adopted. It should not be thought that this number has been derived rigorously – it is as good as we can get, for as much as we want to spend. There is one detail here – the prototype antenna vendors were not happy to supply an antenna that could "blind point" (point to any direction with only the aid of a global pointing model) to 0.6 arcseconds. They will instead provide an antenna that blind points to 2.0 arcseconds, and then requires a periodic local pointing model update to reach 0.6 arcseconds. In terms of the calibration, we must then make sure that we can derive the quantities of the global pointing model to allow for the 2.0 arcsecond pointing, and the quantities of the local pointing model (and the frequency of updating it) which allow for the 0.6 arcsecond pointing. If these requirements cannot be met, we are faced with the prospect of developing active metrology.

While we do not divide up the 0.6 arcsec pointing error specification into various systematic and random terms here (but see Table 3.4.3-2 of the ALMA Antenna RFP), we note that the effects of any pointing error budget with various systematic and random terms could be translated to an estimated image quality, given the image (see, e.g., Pety et al. 2001b). We also note here that random errors have less effect than systematic errors, while pointing out that it may be possible to correct for some of the effects of the systematic errors. With only minor exceptions, pointing calibration must be performed prior to astronomical observations or the data are irretrievably corrupted (but see the note on pointing selfcalibration below). This also means that we cannot generally interpolate pointing solutions backwards in time. This makes pointing calibration critical.

2.1.2 Surface Setting and Primary Beam

In order to maximize the forward gain of the antenna, the surface errors must be minimized. In addition, surface errors can be the limiting factor in mosaicing of large complex sources. Cornwell et al. (1993) derived a requirement of $\lambda/40$ for the surface accuracy, driven by obtaining a fidelity of 100 on large (mosaicing necessary) complex images. For the MMA at 350 GHz, this turned into a surface rms of 22 μm. More detailed analysis of the MMA case showed that 25 μm was probably sufficient for frequencies up to 650 GHz (Holdaway 1992). This surface error was considered reasonable given the 8m diameter of the MMA antennas (see, e.g., Lamb 1993). Even given the move to 12m diameter antennas for ALMA, it was believed that this surface error was achievable (this after discussions with the prototype antenna vendors). The antenna RFPs therefore contained specifications of 25 μm with goals of 20 μm. This, like the pointing requirement, is a matter of getting as good as we can get for the price we want to pay (if we thought it were possible and we could afford it, we might ask for true $\lambda/40$ at 950 GHz, which would be ~8 μm). The difficulty is then in making measurements and adjustments to the surface panels to meet that requirement.

For accurate mosaicing, it is also necessary that the antenna voltage response be precisely known on each of the elements of the array. Cornwell et al. (1993) determine that the primary beam must be known to 6%, in order to reach image fidelity of 100 (in their definition), and in order that errors from this be smaller than those from pointing. We adopt 6% as the requirement on this calibration measurement, and further stipulate that it applies out to the point in the beam where the response is 10% of the peak. Note that this is the power pattern – the requirement on knowledge of the voltage pattern is 3%.

2.1.3 Subreflector and Feed Positioning

Having the subreflector or feeds positioned incorrectly will result in a loss of efficiency and in primary beam abnormalities that may be very hard to correct in mosaicing. Accounting only for the loss of efficiency, and requiring that the loss is < 1%, Butler (2003) calculated the allowable subreflector and feed offsets shown in Table 1 for the ALMA antennas. When the antenna design is finalized, this calculation should be repeated.

Table 1. Allowable Feed and Subreflector Position Errors

Type of Error	**Positioning Requirement**	**Requirement @ 650 GHz**
Feed Axial Offset	0.9 λ	280 μm
Feed Lateral Offset	10.0 λ	3.2 mm
Subreflector Axial Offset (Focus)	0.09 λ	28 μm
Subreflector Lateral Offset	0.45 λ	140 μm
Subreflector Rotation (Tilt)	λ/.63m (rad)	1.7 arcminutes

2.1.4 Motion

In order to calibrate the fluctuations in the atmospheric delay, it is necessary to be able to switch to nearby calibrator sources on short timescales. Combining our knowledge of the distribution of such calibrator sources on the sky with what is realistically achievable in terms of slew speed and settle times for a 12-m antenna produced the current specification that the antenna be able to slew 1.5 degrees and settle down to within 3 arcseconds of the new desired pointing location in 1.5 seconds.

In addition, the antenna must be able to slew quickly across relatively large sources to allow for cancellation of atmospheric fluctuations in single dish data (D'Addario 2003). For a small number of antennas, wobbling subreflectors are required for this same purpose (Pety et al. 2001a).

2.2 Amplitude

For the amplitude calibration, a well defined scientific goal can be elucidated and set as the requirement. In numerous meetings and discussions, the scientific community has made clear its desire to reach 1% flux density accuracy (most recently reiterated in the report of the spring 2003 ASAC meeting), which means that we must be able to determine the overall flux density scale (and apply it to the visibilities and total power measurements) to 1%. In addition, the capability of achieving a dynamic range in images of 10000 or higher means that we must track the amplitude fluctuations to better than 1% (Yun et al. 1998). A later study by Moreno & Guilloteau (2002) showed that it is impractical to achieve 1% at the submm wavelengths, and so a requirement of 3% has been adopted for ν>300 GHz.

2.3 Delay

Because ALMA is a connected element interferometer, with correlation happening real-time, the delay must be calculated correctly at the time of observing. If delay is not calculated properly, we will never be able to properly calibrate ALMA data. The delay involves several components, including source position, earth orbit and orientation, atmosphere (both bulk and fluctuating, and both neutral and ionized), station location, antenna structure delay, and electronic delay. All of these components must be properly calculated or measured. This is very important, as the delay is applied to the voltages before correlation, and hence the effect of an incorrect delay cannot be undone in many cases.

There is an additional fluctuating portion of delay due to the atmosphere, antenna, and electronics. Delay fluctuations limit resolution, limit the dynamic range of images, introduce artifacts, and reduce sensitivity by decorrelation. Without effective calibration of these fluctuations, the maximum usable ALMA baseline would generally be about 300 m. Amplitude errors would limit image dynamic range and skew the flux density scale. Therefore, high fidelity imaging of any kind (mosaiced or not) requires accurate delay calibration.

The phase calibration working group report (Woody et al. 1995) considered three cases at 300 GHz: high quality imaging with 8^o total phase errors, median conditions with 19^o deg phase errors, and poor imaging with 48^o phase errors (these phase errors are antenna based – baseline based errors are $\sqrt{2}$ larger). The phase errors have a budget which includes the atmosphere, the antenna, and the electronics. Converted to path lengths, these correspond to 23, 52, and 134 μm, or, again, converted to delays (in time), these correspond to 77, 172, and 446 fs (femtoseconds!). This has been updated based on extension of ALMA to higher frequencies, and a recognition that when the atmosphere is less dry we will not be observing at the higher frequencies. D’Addario et al. (2002), further recognizing that the delay can have both a fluctuating and systematic offset portion, specified that the requirements on the delay setting and measurement were:

error (fs)	τ_a	τ_s	τ_e	total
systematic	8.4	4.8	6.9	11.9
random	38.5	22.2	31.4	54.5

where τ_a is the atmospheric delay, τ_s is the antenna structure specific delay, and τ_e is the electronic delay.

We will now refine slightly these delay measurement and setting requirements. We define the total system delay at the i^{th} antenna as:

$$\tau_i = \tau_{g_i} + \tau_{a_i} + \tau_{s_i} + \tau_{e_i} + \tau_{f_i} \tag{n}$$

where τ_g is the geometric delay, τ_a,τ_s, and τ_e are as above, and τ_f is a fudge factor subsuming the errors in the other quantities. We specify the following requirements on the different parts of the delay:

error (fs)	τ_g	τ_a	τ_s	τ_e	total
systematic	5	10	7	7	15
random	0	35*(1.25 + PWV)	50	30	$\sqrt{5300 + 3100\,\mathrm{PWV} + 1200\,\mathrm{PWV}^2}$

Where PWV is the precipitable water vapor in mm. We now discuss in more detail some of these requirements and requirements on other related measurement quantities.

2.3.1 Source Location

In order to accurately calculate the geometric delay, it is necessary to know where the desired source (or interferometric phase pointing center) is. We require that the phase center can be pointed to a fraction of a synthesized beam width for the most spread out configuration. Given 14 km baselines (the longest baselines are 18 km, but the equivalent baseline length to give the resolution of the most spread out configuration is ~15 km – Otárola & Holdaway 2002), and 950 GHz frequency, the synthesized beam FWHM is ~4 milliarcseconds. It seems reasonable to require that the phase center be pointed with an accuracy of 0.5 milliarcsecond. In order to do this, we propose that only J2000.0 coordinates be used, with proper precession and nutation and timing (Sovers et al. 1998). An additional correction that will be necessary is the relativistic bending of the radio waves in the gravitational potential of the sun (Sovers et al. 1998; Wade 1976). The use of CALC (Ryan & Ma 1979) or a similar package to calculate the geometric delay will automatically include this correction.

2.3.2 Atmospheric Delay

The Earth's atmosphere provides an additional component of delay between the two antennas. There are two possible parts of the atmospheric delay term that will have to be accounted for, the neutral atmosphere (mostly the troposphere), and the ionosphere.

2.3.2.1 Neutral Atmosphere

For the most part, the differential delay between the two antennas due to the non-fluctuating part of the neutral atmosphere is the same, and hence can be ignored. However, on the longest baselines, the difference in elevation between two antennas can be of the order of 10 arcminutes, resulting in a differential delay that is significant (at 10 degrees elevation, the difference can be as large as 400 picoseconds). The bulk atmospheric delay must therefore be calculated separately for each antenna.

In addition, there is a fluctuating part of the delay, caused by fluctuations of water vapor in the troposphere. The calibration of this fluctuating tropospheric delay is often called "phase calibration", but we do not use that term here. The current specification on the delay correction for the fluctuating atmospheric contribution for the WVR systems is: 10 * (1 + PWV) μm, where PWV is the precipitable water vapor column in mm (Hills & Richer 2000; Hills et al. 2001). Note that this is in disagreement with the value adopted in Chapter 7 of the Project Book, which lists the allowable atmospheric phase path error as 38.5 fs, regardless of conditions, i.e., leaving off the PWV component (see Table 7.4).

There is a further fluctuation due to the fluctuations in density and temperature of the dry (non-water) part of the neutral atmosphere. It is likely that this term is small, but we do not know this definitively yet, and so allow for some small additional dry component fluctuations in the overall error budget for atmospheric fluctuations.

2.3.2.2 Ionosphere

An effect that is often ignored in millimeter interferometry is the delay induced by the ionosphere. Given the extreme accuracy required by ALMA, we may not be able to afford to completely ignore it. The zenith ionospheric delay at 8.5 GHz is typically of order 1 ns. Since the ionospheric delay goes like ν^{-2}, this would result in about 100 ps at 30 GHz. It is not the total delay we are interested in, but the differential delay on 18 km baselines. While this should be smaller than the total delay, it still can be significant (experience at the VLA tells us this), and note that the delay is larger at lower elevations (see the discussion in Sovers et al. 1998). A recent study by Hales et al. (2003) concluded that fast phase fluctuations seen in the site phase interferometer are due to ionospheric fluctuations. The magnitude of these fluctuations can be as large as 30^o rms at 11 GHz. Scaling this to 30 GHz gives fluctuations of order 5^o. Not overly large, but beginning to be of some concern. The scaling with baseline length and elevation is not well constrained. So, we do not have enough information at present to tell whether we will definitely have to monitor the ionosphere, but indications are that we might. For this reason, we allocate a small additional amount to the error budget for atmospheric fluctuations to account for this.

2.3.3 Antenna Location

The relative positions of the antennas must be determined accurately so the geometrical delay can be correctly calculated and supplied to the correlator. Residual delays due to incorrect antenna locations will result in phase errors which change across the observing band and differential phase errors between two different sources on the sky (for example, a calibrator and target source). For ALMA, with an 8 GHz bandwidth, a reasonable limit of 1/3 radian phase difference across the band requires a baseline accuracy of about 1 mm. The differential phase error between two sources on the sky can be written as (Sramek 1981; Wright 2002):

$$\Delta\phi=\Delta x\Delta\theta \qquad (1)$$

where $\Delta\varphi$ is the phase error in turns, Δx is the baseline error in wavelengths, and $\Delta\theta$ is the separation of the sources on the sky in radians. We take 10^o as the maximum allowable phase error at the highest observing frequency (950 GHz). We are more conservative with baseline errors than with atmospheric errors because they will be partially systematic. For a source separation of 5^o, if the maximum error is constrained to be 10^o at 950 GHz (equivalent to 9 μm of path), then the error in the baseline must be less than about 90 μm. This means that the error on an individual antenna location determination must be less than about 65 μm, if the errors add in quadrature (not true for any given baseline, but averaged over the array, this is a reasonable assumption). Note that the antenna location determination is needed as a precalibration, as it is used to determine the geometric delays applied to the antenna signals before correlation, but it can also be used in postcalibration for the baseline determination (u,v,w can be recalculated based on better antenna location [baseline] calibration).

Efforts will be required to properly understand how the effective antenna location changes with time. On the longest baselines, earth tides and polar motion will have significant effects, and we will need to borrow the techniques of VLBI for antenna location determination (Clark 1973;

Sovers et al. 1998). Once again, the use of CALC (Ryan & Ma 1979) or a similar package for the calculation of the geometric delay model will automatically include these terms.

Note that the ability to determine the antenna location to a particular accuracy also provides a requirement for the accuracy of timing across the array. The required timing accuracy from this one consideration is:

$$\Delta t = \frac{\Delta x}{\omega_e B_{max}} \tag{2}$$

where ω_e is the rotation rate of the Earth, and B_{max} is the maximum baseline. For 65 μm accuracy over 20 km baselines this implies a timekeeping accuracy of ~ 45 μs. This is easily met by the current design (D'Addario 2000).

2.3.4 Antenna Structure Delay

There is an additional component of delay introduced by the structure of the antenna itself, and its change with azimuth and elevation. These can be thought of as modifications to the antenna location, in many ways. There is a full discussion of this in the U.S. Antenna RFP, section 3.4.5.

Repeatable (and thus perhaps predictable) parts of this delay include the change of main reflector shape with elevation or azimuth (e.g., nodal point changes with tilt), subreflector position with elevation (see, e.g., what is done with axial displacement ["focus"] as a function of elevation at the DSN antennas in Jacobs & Rius 1990) axis non-intersection (Butler 2003; Wright 2002; Sovers et al. 1998; Wade 1974), the illumination offset (Holdaway 2001a), bearing runout, and bearing alignment. The specification on this component in the Antenna RFP is 20 μm, equivalent to 67 fs. However, most of this repeatable component can be accurately predicted, if measured properly, so we do not allocate 67 fs to this term. We estimate that only about one tenth of the repeatable residual delay will result in true systematic delay offsets, so allocate 7 fs to this term.

Nonrepeatable (and thus probably not predictable) parts of this delay include the change of main reflector shape, subreflector position, or feed position with temperature changes (causing thermal deformation of the structure) or wind or other acceleration forces, and bearing nonrepeatability. In fact, some of these nonrepeatable errors may be predictable, but we cannot count on it. The total of all of these effects will likely only be answered by experimentation over time with the final ALMA antennas. In order to be somewhat conservative, we adopt the Antenna RFP value for the nonrepeatable residual delay as the allowable random antenna structure delay error. This value is 15 μm, or 50 fs. Note that both this and the repeatable delay error are specified for motions within a 2^{o} solid angle on the sky, and the nonrepeatable portion is specified for timescales less than 3 minutes.

2.3.5 Electronics Delay

Finally, there is a component of delay introduced because of the electronics between the feeds on the antennas and the samplers. This will need to be measured for each antenna, receiver cartridge, and polarization. The determination of this delay should be accurate to a fraction of the other error and offset terms, if possible. We adopt the numbers recommended in Chapter 7 of the Project Book.

2.4 Bandpass

ALMA will always operate as a spectral line instrument, in the sense that multiple frequency channels will always be present in the product of the correlator. As such, knowledge and calibration of the shape of the frequency response of the instrument and atmosphere is very important, and can limit not only the spectral dynamic range (in the case of observations of spectral lines), but also the image fidelity (in the case of continuum imaging, since bandpass variations are effectively like amplitude variations if not calibrated properly).

The required spectral dynamic range is driven by a few types of experiments that need to observe weak spectral transitions against a strong continuum. These include observations of rare isotopes (extremely optically thin) against quasars (potentially many Janskys of flux density), and HDO emission (weak emission feature on limbs, even weaker in absorption against the disk) in the atmosphere of Mars (extremely strong continuum). In these cases, a dynamic range of 10000 to 1 is required. For most observations, however, a dynamic range of 1000 to 1 is sufficient.

2.5 Polarization

Because measurements of the polarization of sources yield intrinsic information on the physical parameters of those sources, it is a requirement that ALMA be able to produce visibilities simultaneously in all four Stokes parameters (or some equivalent, from which they can be derived) when asked for by the observer. Measurement of the polarization morphology of isolated protostars, circumstellar disks, or or protostellar clusters requires that the polarization percentage be measured to 0.1% in amplitude, and that the position angle of the linear polarization be measured to an accuracy of 6^o. With these accuracies, magnetic field geometries in protostellar outflows can also be measured, providing a test of the mechanism of generating these outflows.

2.6 Sideband Gain Ratio

Since some receivers will be double sideband (DSB), they will have emission from an unwanted part of the spectrum in the single-dish signal. In order to reach the overall 1/3% amplitude calibration requirement in the single-dish data, it will be required to determine the sideband gain ratio to a small fraction of this – 0.1%.

2.7 Summary

Pointing	2.0 arcseconds blind; 0.6 arcseconds offset
Primary Beam	6% in power out to 10% point
Feed Setting	280 μm vertical; 3.2 mm lateral
Subreflector Setting	28 μm vertical; 140 μm lateral; 1.7 arcminutes rotational
Antenna Motion	1.5 degrees in 1.5 seconds with settle to 3 arcseconds; XXX more here

Amplitude	1% for $\nu<300$ GHz; 3% for $\nu>300$ GHz
Antenna Location	65 µm
Geometric Delay	5 fs systematic
Atmospheric Delay	10 fs systematic; 40 (1.25 + PWV) fs fluctuating
Antenna Delay	7 fs systematic; 50 fs fluctuating
Electronic Delay	7 fs systematic; 30 fs fluctuating
Bandpass	1000:1 in most cases; 10000:1 in a few select cases
Polarization	0.1% in amplitude; 6^{o} in position angle
Sideband Gain Ratio	0.1%

3 Precalibrations

Precalibrations are those calibrations to determine quantities necessary to make adjustments to the instrument which affect the signal prior to correlation. In almost all cases, these adjustments are irreversible. The quantities may vary slowly, in which case they will generally be determined in infrequent special calibration observations, or they may vary more quickly, in which case they will need to be determined by measurements during the observation.

3.1 Antenna Characteristics

3.1.1 Pointing

The antenna pointing will be affected by several slowly varying terms such as systematic imperfections of the antenna and the pad, gravitational forces, and thermal loading from the Sun. Depending upon the strategy of the astronomical observations, much of the effect of these slowly varying quantities can be removed by frequent offset pointing observations. Some active corrections might also be incorporated into the antenna design. Thermistors have been employed on existing antennas to monitor solar thermal response (Lamb & Forster 1993; Lamb & Woody 1998a; Bayley et al. 1994), and their use, along with tiltmeters and other metrology devices (lasers, e.g.) in the ALMA antennas has been proposed and investigated (Anderson 1999; Plathner et al. 1999; Woody & Lamb 1999; Kingsley et al. 1999; Plathner 1999; Lugten et al. 1999; Lugten 1998; Payne 1997). The antenna contract stipulates that the contractor supply any extra metrology devices deemed necessary to meet the pointing specification, and allow for the addition of other metrology devices by the ALMA project, including laser/quadrant detectors, tiltmeters, temperature probes, and laser/retroreflector systems.

In addition to these slowly varying systematic pointing errors, there will also be highly random pointing errors caused by wind loading and dynamic refraction (often called "anomalous refraction", but we dislike that term, and adopt the one proposed by Pety et al. 2001b [note that Barry Clark once proposed that this be called "nomalous refraction"]). Finally, at some level there will be a limit to the mechanical repeatability of the antenna pointing. At this level, we are left with completely random pointing errors which cannot be calibrated. If these purely random errors are too large, they will spoil the imaging characteristics of ALMA and will not be correctable. If they are small to moderate in magnitude (i.e., < 0.5 arcsec), we can tolerate them quite well as these random errors are the least damaging of any pointing errors.

Our goal in determining the pointing parameters of the antenna is to remove all of the systematic pointing errors (as best we can), leaving only the small purely random errors. We do this in three ways: through use of a good global pointing model; by using local offset pointing determination; and by determining refraction (both static and possibly dynamic) offsets.

3.1.1.1 Global Pointing Model

The first step in removing the systematic pointing errors is to determine the systematic imperfections of the antenna and pad and the effects of gravity. Most radio telescopes periodically undergo a pointing routine which samples the sky with pointing measurements on about a hundred astronomical sources taken at night to minimize thermal and wind pointing errors. ALMA will take about 60 minutes to perform 100 pointing calibration observations across the sky (each one consisting of one or more pointing patterns around a strong quasar – see below). The results of these pointing measurements (beam offsets in azimuth and elevation) are then used to fit about 10 parameters in a pointing equation which accounts for various physical terms, such as misalignment of optical axes or four fold sag due to the antenna base being supported at four locations. Some experimenting will go into determining the optimal form for the ALMA pointing equation (Mangum 2001). In particular, the precision of the refraction correction is essential (see below). We note here that the pointing calibration observations (the pointing patterns) can be performed either in total power or interferometric mode, but the interferometric mode is significantly more sensitive, so we anticipate using it in almost all pointing calibration observations.

One point of note is that the determination of the axis non-intersection term is difficult unless the antennas are commanded to go "over the top" during the pointing parameter determination, at least on some sources (Butler 2003). This is a requirement for the ALMA antennas, and it will need to be utilized for the determination of this parameter.

ALMA will probably rely heavily upon a lower frequency (probably 90 GHz or below) system for determining the pointing model. The wider beam at this low frequency and the high sensitivity and denser grid of bright astronomical sources will facilitate pointing measurements, even after a reconfiguration. However, precise pointing offsets among the different frequencies will also need to be determined (often referred to as the "collimation offsets").

One question is how many antennas will be required to do an antenna pointing model determination? Given the current operational model of frequent antenna moves (roughly one per day), this question is quite important. After an antenna is moved, it will be required to determine its pointing model parameters, and this will require some number of other antennas to be used for this. In practice, some number around 3-6 are used in current arrays, but this number remains to be determined.

3.1.1.2 Local Offset Pointing

The requirement for the blind pointing after application of the pointing model is 2 arcseconds rms. To achieve the precision relative pointing specification of 0.6 arcseconds, frequent local offset pointing calibration will be required, to remove local deviations from the pointing model, or systematic slowly varying effects due to wind and/or thermal gradients. This type of local offset pointing has been used very successfully at many observatories for many years now (for the example of the VLA, see Rupen 1997).

Holdaway (1996) and Lucas (1997) have both demonstrated the feasibility of performing this kind of local offset pointing with ALMA – adequate SNR can be obtained with sufficient speed. The minimum calibrator flux density, and hence the typical minimum distance to a pointing calibrator for pointing calibration, is a function of both the collecting area and the number of elements in the array.

A key question concerning the efficiency of these offset pointing calibrations is: what are the differential pointing errors as a function of distance between calibration and target sources? This depends upon the direction of wind, the sun angle, the flux density of the calibration source, etc…, and probably needs to be answered experimentally. Also, we must understand the stability of the pointing offsets among the different frequency bands, as we would prefer to do the pointing calibration at 90 GHz, but if the collimation offsets are changing this will not be possible. This understanding will only come after the antennas are complete and we can test for these particular issues.

It might be possible to minimize the differential pointing errors by performing a local average on a few nearby strong (0.5 Jy) quasars (Moreno & Guilloteau 2002). Simulations should be performed to check the impact of dynamic refraction in such a mode, and to determine the optimal strategy for this kind of "local pointing model" (how often to update, how many pointing calibrators, position relative to target source, necessary flux density, etc…).

Both the pointing model solution and the offset interferometric pointing will require an extensive, up-to-date catalog of pointing calibrator sources, and the observing schedule program should allow for automation of the choice of a pointing calibrator and the pointing calibration strategy.

Measurements of the wind at the site (see, e.g., Sakamoto et al. 2000) indicate that a great deal of the power in the wind is often in a constant speed and direction, and this wind at constant speed and direction will often be corrected by the local offset pointing calibrations. In that case, it is only the gusts about this mean speed and direction which will result in differential pointing errors between the calibration and the target source.

We do not know currently how often we will need to perform offset pointing. Experience on the VLA is that under benign conditions every hour or so is sufficient. Around sunrise or sunset, it must be done every 15 minutes or so. Only experience with the real ALMA will be the guide as to how often the offset pointing must be done on the array.

3.1.1.3 Pointing Pattern

When observing the calibration sources during either a global or local offset pointing calibration observation, a number of positions in the beam must be observed – these comprise the "pointing pattern." There have been many implemented pointing patterns at radio telescopes. The most common is the "5-point", where a point at the center, and at the 4 half-power points (N and S in elevation, E and W in azimuth), is taken. In addition, one or more "off" positions might be measured (but only if observing in total power [single dish] mode). A variant of this is the one used on VLBA antennas, called the "Craig Cross" (named after Craig Walker, its inventor). This is a 10-point pattern which attempts to make a better baseline subtraction, especially in the elevation direction – useful for total power observing. The IRAM 30-m has used scans in azimuth and elevation instead of discrete pointings (Greve et al. 1996). The OVRO pointing pattern is a series of rotating triangles at the half-power point, repeated until the pointing error reaches a specified level (Scott 1993). Scanning circles have been employed at Arecibo (Olmi & & Davis 1998). We do not know which of these patterns will be employed for ALMA at this time.

3.1.1.4 Refraction

The bending of the radio waves in the atmosphere of the Earth (refraction) must be accounted for when calculating the instantaneous pointing direction to a given source. This is an explicit additive term to the elevation offset in the pointing determination. The bulk of the refraction is due to the dry component of the atmosphere, and is relatively well determined given surface measurements of temperature and pressure. We follow the recommendation of Pety et al. (2001b) and call this the "static refraction." There is an additional rapidly fluctuating refraction component which is due to the turbulent water vapor above each antenna. Again, we follow the recommendation of Pety et al. (2001b) and call this the "dynamic refraction."

3.1.1.4.1 Static Refraction

This is the refraction of the radio wave due to the dry and non-fluctuating wet atmosphere. It causes a deflection in the direction of elevation, i.e., at low elevations, a source appears to be at a lower elevation than it really is. There have been many studies of this effect for radio telescopes (see the discussion in Mangum 2001 and references therein). Note that it requires an accurate measurement of the surface temperature, pressure, and humidity. Mangum (2001) give the following measurement requirements: ambient pressure: 0.5 mb; ambient temperature: 0.1 C; relative humidity: 0.5%; wind speed: 0.5 m/s; wind direction: 5^o. Note also that the difference in elevation at the different antenna locations can provide a relatively large offset in the refraction correction, so that correction will have to be calculated for each of the antennas separately. For example, two antennas separated by 16 km will have an elevation pointing difference of roughly 10 arcminutes, and at 10 degrees elevation, this results in a refraction correction difference of about 4 arcseconds (out of 190 arcseconds total refraction correction). Since this is larger than even the blind pointing requirement, it will have to be accounted for.

3.1.1.4.2 Dynamic Refraction

This fluctuation in the pointing is a natural consequence of the fluctuation water vapor above the antenna. It is akin to the fluctuations in the dry atmosphere which cause so-called "Angle of Arrival" fluctuations on optical/IR telescopes, because the bulk of the effect of the fluctuating refraction is a simple change in the effective direction between the propagating plane wave and the antenna optical axis.

The effect of dynamic refraction on the pointing will usually not be a severe problem, but will sometimes limit the pointing (Holdaway & Woody 1998a; Lamb & Woody1998b; Holdaway 1997b; Butler 1997). Since the refractive pointing is random on time scales of the antenna crossing time of the atmosphere (~ 1 s), we merely need to have many statistically different short measurements for each of the points of a pointing pattern to ensure that we are not applying an erroneous pointing position when we collect data on the target source. Note that this averaging process contradicts the idea of performing "fast" pointing calibration every 30 min or so. Also, it is in apparent contradiction with Lucas (1997) which shows a 1:1 correlation between the rms pointing error and the anomalous refraction, although the pointing was measured on 1 min timescales.

There is the possibility of correcting for the dynamic refraction with additional devices, or at least measuring what the offset due to anomalous refraction is (Lamb & Woody 1998b; Hills & Richer 2000). These are still under study.

3.1.1.5 Pointing with Optical/IR Telescope

The ALMA prototype antennas, and perhaps the production antennas, will be outfitted to perform infrared or optical offset pointing, with a small telescope looking through a hole in the antenna primary surface (Mangum 2000). Offset interferometric pointing should work well enough. However, optical/IR pointing will increase the overall efficiency of the instrument, may improve the pointing of the antennas and may help characterize mount components in the pointing equation. The infrared pointing will be largely immune to dynamic refraction effects (since it is such a small aperture, normal optical/IR angle-of-arrival [AOA] fluctuations will be minimal). We have not determined whether an optical/IR pointing telescope will be needed on all ALMA antennas yet, this is still under study and discussion. It is clear that such a telescope will be of great benefit during the verification and commissioning of the antennas, however. At the very least each antenna should allow for the mounting of such a telescope, even if there is not one present on all antennas at all times.

3.1.2 Antenna Surface and Primary Beam

In order to maximize the forward gain of the antennas, and to make the beams as well behaved as possible, it is desirable to measure the deviations of the surface from the optimal surface, and to correct for them (by adjusting the surface panels). We will use the technique of interferometric holography to do this (Scott & Ryle 1977).

The requirement on the holography raster size can be derived from the fact that there should be roughly four resolution elements across the smallest panel dimension (Bennet & Godwin 1977). For the current prototype designs, this is roughly 0.5 meters. This implies that a 128 x 128 raster will be required for holography (if a power of two is required – otherwise it would require a 96 x 96 raster). If holography is done at 90 GHz, and if the oversampling factor is 10%, then this is a total angular size of roughly 1.8^{o}. Given that we would like to do holographic measurements on a timescale that is short relative to any thermal deformations, this implies sample times at each raster point of about 0.4 seconds (2 hours per raster, with a boresight every 15 seconds).

The surface measurement error should be at least a factor of two less than the desired surface accuracy. Given a 20 μm surface accuracy goal (see the antenna RFPs), this implies a surface measurement accuracy of at least 10 μm.

The rms surface measurement error arising from measurement at wavelength λ of a source with SNR S over an N x N raster can be written (Butler 1999):

$$\varepsilon \sim \frac{\lambda N}{5 \pi S} .$$

For a frequency of 90 GHz, a required rms of 10 μm, and a raster size of N = 128, this gives a required SNR of S > 2700. Given a sample dwell time of 0.4 seconds, this implies a source flux density of F > 45 Jy (assuming a single baseline and 17 mJy noise in 0.4 seconds on that single baseline). The required flux density can be reduced if many reference antennas are used, but at an operational penalty (these antennas are not available for normal observing). There are not very many sources with flux density of this magnitude – planets are an exception, but are often resolved (and hence require a correction based on that fact, which increases the uncertainty since the sky brightness distribution is never perfectly known). The required SNR is decreased with shorter wavelength, but the noise goes up in almost the same proportion. We may, therefore, be required to have a transmitting beacon on a nearby mountain top (Cerro Chascon or Cerro Chajnantor, perhaps). If this is the case, care must be taken to properly apply the second order (and possibly higher) near-field corrections.

The rms surface measurement error on antennas with dish diameter D arising from random pointing errors θ_{rms} is (Butler 1999):

$$\varepsilon \sim \frac{\theta_{rms} D}{12} .$$

Given 0.6″ rms pointing errors, this yields a measurement rms of 3 μm. Not unreasonable, but it must be kept in mind when calculating the entire measurement error budget.

Amplitude fluctuations (which can come from electronics or atmospheric scintillation) of magnitude f_a cause an rms measurement error of (Butler 1999):

$$\varepsilon \sim \frac{\lambda f_a}{15 \pi} .$$

For an amplitude fluctuation of 10^{-3}, this results in a measurement error of 0.1 μm, which should be negligible in the error budget. Note that this assumes that we track the amplitude scintillations to that level – if left uncorrected, the level might be larger than this by an order of magnitude (still a small part of the entire budget).

Phase fluctuations of magnitude ϕ_{rms} (or, equivalently, path length errors of Δ_{rms}) result in a measurement error of (Butler 1999):

$$\varepsilon \sim \frac{\lambda\ \phi_{rms}}{12\ \pi} \sim \frac{\Delta_{rms}}{6} .$$

We expect to correct the phase fluctuations to $\Delta_{rms} < \sqrt{490 + 250\ \mathrm{PWV} + 100\ \mathrm{PWV}^2}$ μm. For PWV of 1 mm, this is $\Delta_{rms} \sim 29$ μm, or a measurement error of ~ 5 μm. We will have to use good weather conditions when performing holography measurements.

There are other issues to be concerned with – bulk reflector shape change (with elevation, mostly), when using astronomical sources; ground scatter if using a beacon; unmodelled EM clutter from the structure itself, etc…, but most of them are of less concern than the above points.

Holography measurements also provide us with a measurement of the primary beam shape of the antennas. However, a holography raster which is optimized to measure the small-scale deviations in the surface (with large spatial extent on the sky) is not optimal for measuring the larger scale deformations which affect more seriously the inner part of the primary beam – that part which is most important for mosaicing. In order to measure the inner part of the beam accurately, a different raster should be used, the extent of which is only out to a few FHWM or so, and heavily oversampled.

Note that the primary beam patterns must be known as a function of both frequency and polarization. The measurements should be done to cover this parameter space. In addition, it may be desirable to have information on how the primary beam pattern changes as a function of elevation or other conditions. The full parameter space may be too wide to sample fully, but we should gain some experience during early operations on how finely to sample it.

Holography measurements also provide us with a measurement of subreflector and feed positioning errors.

3.1.3 Subreflector and Feed Positioning

Subreflector and feed positioning errors reduce the on-axis gain of the antennas, and distort the primary beam shape in bad ways. Errors in feed positioning are most easily measured by doing holography. However, a much more efficient way of getting the subreflector positioning errors is to do scans while looking at a strong source. Since we will have a subreflector positioning mechanism which can be commanded by software, it is easy to do not only the traditional so-called "focus curves" (which measure the optimal axial position of the subreflector), but in addition to do "tilt curves" (which measure the optimal rotational angle of the subreflector) and "offset curves" (which measure the optimal lateral position of the subreflector).

3.2 Amplitude

The amplitude calibration contains no required component of precalibration. If quantities are available at the time of observation, then the visibilities and single dish powers can be immediately put onto a realistic amplitude scale, but this is not necessary, and it can all be done after the fact (in postcalibration).

3.3 Delay

3.3.1 Source Location and Geometric Delay

As mentioned above in the requirements section, we must be able to accurately specify source locations (or, phase pointing locations) accurately. In order to do this, we should use only J2000.0 coordinates, with proper precession and nutation. In addition, we should account for all modifiers of geometric delay that we can reasonably include in the model. We recommend that a package similar to CALC (Ryan & Ma 1979) be used to calculate geometric delay, which should include polar motion, earth tides, and other effects. We will have to obtain Earth Orientation Parameters from some source – we recommend the USNO for this purpose (similar to what is done with VLBA, VLA, and ATCA currently). These parameters are changed on roughly month long timescales, and we should implement a scheme to automatically incorporate them into the geometric delay model if possible (as ATCA has done).

3.3.2 Antenna Location and Structure

In order to accurately calculate the geometric delay before correlating the voltages from two antennas, their locations must be known. The process of locating these antenna positions is often called "baseline calibration" because it allows for accurate specification of the baseline as well, but we prefer to simply call it "antenna location calibration" since that more accurately describes what we are measuring/calibrating.

The antenna locations may be measured by determining the delay on each baseline for on order of a hundred observations of point sources sampling the entire sky. Individual delays can be fit across the spectrum, as in VLBI. The complete set of delays is used to solve for the three dimensional locations of all antennas relative to a reference antenna (or any reference location). Fomalont & Perley (1999) and Thompson et al. (2001) give good general overviews of this measurement. Wright (2002) has given a good description of this process on the BIMA array. Sovers et al. (1998) give a description of the technique used in VLBI.

The observing strategy is similar to that for the pointing model determination, and should take about an hour to complete. Signal to noise is not an issue for 65 μm accuracy, and the 1 hour time scale is set more by the minimum time to sample many sources around the sky. Atmospheric phase fluctuations may affect the baseline delays, so ideally the observing conditions should be excellent. In poor conditions, the delays can probably still be determined based on the statistics of many differential measurements, as the atmosphere should tend towards a zero mean in differential measurements. A timescale longer than 1 hour is required for baseline measurements to reach the outer scale of the atmosphere and sample several atmospheric screen patterns. Local rotation of the baseline plane is possible otherwise. Experience in the early phases of operations will be required.

In order to reduce errors during antenna location determination, only sources with well determined locations should be used (Feissel-Vernier 2003; Johnston & de Vegt 1999). In addition, since source structure will introduce further uncertainty in the antenna location determination (Sovers et al. 1998), sources with very simple structure should be chosen, otherwise we must use a more complicated self-calibration procedure, obtaining both the structure and the location.

Of some concern is the time scale over which we can expect the antenna positions to remain fixed to within 65 μm. Permafrost has been reported on the ALMA site (see, e.g., Snyder et al. 2000), which enables an entire class of soil movements. We can probably expect some amount of soil creep, especially after earthquakes (see Otárola et al. 2002). We will gain experience concerning the frequency of baseline calibration, or the possibilities to correct for some of the changes using on site measurements, once ALMA begins to be operational at Chajnantor.

One question is how many antennas will be required to do an antenna location determination? Given the current operational model of frequent antenna moves (roughly one per day), this question is quite important. After an antenna is moved, it will be required to determine its location (or, equivalently, its delay), and this will require some number of other antennas to be used for this. In theory, this could be done with a single antenna with accurately known position, but in practice, several are always used. But whether this means 3 or 6 or some other number remains to be determined. We want as many as possible to increase the SNR during the observations, but as few as possible because it takes functional antennas away from normal observing.

3.3.2.1 Subreflector Position Change with Elevation and Other Conditions

The positioning of the subreflector (axial, lateral, and rotational) which maximizes antenna gain will change as a function of elevation and ambient conditions (notably temperature). There are two possibilities for determining this optimal position and its effect on the delay. First, we should measure, very carefully, for each antenna, what the optimal position is as a function of elevation. If this optimal position is stable enough over time, with given conditions, then we may simply be able to use a look-up table (as a function of, say, elevation and temperature) to find this optimal position and the delay term. If not, then we must go to the second possibility, which is much more time consuming, but might be required in the end – periodic measurements of the optimal position may be required during observation.

3.3.2.2 Axis Non-intersection Determination

The fact that the elevation and azimuth axes do not intersect exactly (nor are the angles between them 90°) provides an apparent offset of the antenna position (Butler 2003; Wade 1974; Wright 2002). This axis non-intersection term can be determined during pointing runs as long as "over the top" pointing is allowed on the antennas (Wade 1974), which it will be for the ALMA antennas. This will be done as part of the normal pointing model determination runs for antennas, but may not have to be done every time, as it should vary slowly (on the VLA, for instance, it is only done once every 16 months or so). Only experience will tell us how often we need to remeasure and update these quantities.

3.3.2.3 Illumination Offset

One additional modification to the antenna position is a result of the fact that the receivers are not illuminated perfectly on-axis. This effect, and its measurement and correction are described in Holdaway (2001a). Although we should attempt to measure this and correct for it at the time of observation, much of the introduced error can be corrected in postcalibration by adjusting the effective (u,v,w) coordinates of the visibilities and modifying the primary beam pattern.

3.3.3 Atmospheric Delay

3.3.3.1 Neutral Delay

Although the differential neutral atmospheric delay from antenna to antenna can be determined by postcalibration, we would like to get it as close as possible to begin with, before correlation. If the CALC (Ryan and Ma 1979) package is used to calculate the delays, it includes the provision for calculating neutral atmospheric delay, however it is widely accepted that the atmospheric model used therein is not very accurate. It would be much better to use the ATM model (Pardo et al. 2001) to calculate the expected neutral atmospheric delay. This model will require some inputs – measurements where available (surface meteorological parameters, for example), or estimates where not (atmospheric profiles, if not actually measured). The surface meteorological quantity measurement requirements given in section 3.1.1.4.1 above should suffice for this purpose.

3.3.3.2 Ionospheric Delay

Similar to the neutral delay, we would like an estimate of the bulk ionospheric delay above each antenna if it is available. This would require some knowledge of the total electron content above each antenna, which is probably not feasible. We therefore do not expect to input this as a normal part of the precalibration quantities (though the provision should be there in the software).

3.3.4 Electronic Delay

In addition to the geometric delay introduced from other factors, there is an additional component introduced from the path of the signal through the electronics for each antenna, frontend, and backend. This electronic delay must be determined each time an antenna is moved to a new pad, and if it varies as a function of time on a pad, must be tracked also. Once the antenna positions are well-determined, it is relatively easy to measure the electronic delay for an antenna because it introduces a phase slope across the spectrum (Thompson et al. 2001). All that is necessary is the observation of a strong source, and measurement of this phase slope (it could also be done by looking at the amplitude loss from the delay error but looking at the phase slope is much more accurate). The delay will have to be determined separately for each receiver band,

and also for the cross polarizations. Because of this, the observed source should have strong linear polarization (it is unimportant at what angle, it just needs to be strong).

Only time and experiment will tell us whether the electronic delays remain stable over time. We expect that they should (D'Addario 2000), but should test this in the early days of ALMA to gain practical experience with it.

3.4 Bandpass

There is no component of bandpass calibration that requires a precalibration. While it might be nice to be able to do this calibration real time in order to see realistic spectra, we do not find it compelling enough to cause us to include this in precalibration quantities. If it turns out, in practice, that some parts of the bandpass are very stable (the instrumental response, for example), then it may be possible to measure them infrequently and store them as templates to be used when looking at spectra in real time. We know, however, that there are parts of the bandpass that change with time more frequently (the atmosphere, for example), so we cannot assume that we can precalibrate the bandpass entirely.

3.5 Polarization

There is no component of polarization calibration that requires a precalibration. While it might be nice to be able to do this calibration real time in order to see realistic polarization results, we do not find it compelling enough to cause us to include this in precalibration quantities. If it turns out, in practice, that some parts of the polarization calibration are very stable (the instrumental response, for example), then it may be possible to measure them infrequently and store them as templates to be used when looking at results in real time. We know, however, that there are parts of the polarization calibration that change with time more frequently (the atmosphere, for example), so we cannot assume that we can precalibrate the polarization entirely.

4 Postcalibrations

4.1 Antenna Characteristics

4.1.1 Pointing Selfcalibration

It has been proposed that it might be possible to derive pointing offsets from the data themselves, in mosaiced observations – so-called "pointing selfcalibration" (Wright 1997). While it might work in theory, this technique has never been used in practice. Further, it may not work well at all unless there is at least one bright point-like source within the target area. Once the antenna pointing offsets have been determined, it is simple and not too CPU-expensive to apply the mean array offset as a function of time to mosaic or single pointing data and use that in the imaging. However, if there is significant scatter among the antennas' pointing positions at each time, imaging wide field sources considering the correct pointing data may be prohibitively expensive (Holdaway 1993).

4.1.2 Primary Beam

While we expect to measure the primary beams of the antennas as a precalibration (as part of setting the surface to its optimal shape), it may be true that a better measurement of the primary beam pattern is obtained after observations are completed. In this case, the newest (appropriate) measurement should be used.

4.2 Amplitude

4.2.1 Instrumental Amplitude Calibration

Because of the fast fluctuations in the atmosphere, it is necessary to track the changes in system temperature caused by them. This monitoring must be accurate to at least the level that the amplitudes are desired to be calibrated to in the end, i.e., 1% at frequencies < 300 GHz; 3% at higher frequencies. The currently used technique of tracking these fluctuations, the ambient load chopper wheel method, is accurate to about 5% (see Ulich & Haas 1976; Kutner 1978). Since the requirement is 1 or 3%, we cannot rely on this technique to provide sufficient accuracy. We have investigated two alternatives to this scheme: one in which a dual-load system shines through a hole in the subreflector (Bock et al. 1998; Mangum 2002); and one in which an ambient load shines through a semi-transparent vane into the feeds (Guilloteau 2002). Both of these systems could theoretically reach 1% amplitude tracking accuracy, but initial tests have shown that it will be extremely difficult to reach this level in practice, due to several problems (Bock & Welch 2003; Martin-Pintado et al. 2003). Currently, a scheme in which a wire grid is used instead of a semi-transparent vane, and two ambient loads (of different temperatures), is under investigation (Houde et al. 2001; presentations at ALMA Week 2003). A dielectric beamsplitter has also been suggested (James Lamb's presentation at ALMA Week 2003). Whatever the final system, it is clear that some device where multiple temperatures are allowed to be sampled by the receivers is required (Guilloteau et al. 2003).

With any of these systems, it may be necessary to obtain independent estimates of either the atmospheric opacity, or the atmospheric equivalent radiating temperature (or, even better, the profile of temperature through the atmosphere), or both. At frequencies at which the atmosphere is opaque at low elevation angles and partially transparent at high elevation angles, it will be possible to solve for the sky temperature and the opacity with a sky tip. However, at frequencies at which the atmospheric transmission is excellent, a sky tip will only give the product of the opacity and the sky temperature, and the temperature must be assumed to calculate the opacity. Currently, atmospheric models are not sufficiently accurate to measure the opacity and sky temperature at a partially opaque frequency and accurately estimate the opacity at another frequency. This estimate would have to come either from a model of atmospheric opacity, or from emperical results. The best model currently available is the ATM model (http://www.submm.caltech.edu/~pardo/atm.html). The expected accuracy of the ATM model is about 2%. The Water Vapor Radiometry (WVR) system should provide a measure of emission and absorption at its operational frequency of 183 GHz. Actual measurement of the sky temperature and water vapor profiles via radiosonde are currently underway, along with modeling of the data. Continuous radiosonde monitoring, however, seems unwieldy and expensive. A more cost effective solution would be to float a tethered balloon over the site several times a day. Unfortunately, these systems (tethersondes) are notoriously unreliable and hard to make work under all expected site conditions. The temperature, pressure, and water vapor information would also be useful for the radiometric phase correction schemes.

The best results from atmospheric models may require use of a dedicated sounding device, like an FTS, in conjunction with an atmospheric temperature sensor based on the strong Oxygen lines. This is currently under investigation.

4.2.2 Flux Density Calibration Using Astronomical Sources

In theory, with accurate monitoring of the system temperature, and precise knowledge of other antenna parameters (aperture efficiency, most importantly), an accurate conversion from measured correlated power (or correlation coefficient) can be made. Unfortunately, in practice, it is extremely difficult to know the quantities to enough precision to rely completely on this. Therefore, we are probably going to have to rely on the traditional method of flux density calibration, which relies on astronomical sources of relatively well determined flux density, either *a priori,* or by accurate bootstrapping or measurement at the time of observation.
It is therefore important to find astronomical sources that could serve as flux density standards. The ideal standard should have a predictable flux density as function of time, frequency, and polarization from first principles. Next are standards whose time and frequency dependence can be predicted, but which still require one initial measurement. Finally, sources which are not time variable but still need detailed measurement of their spectrum could be used as flux density standards. Planets are the closest to the first category of standards, but the estimated accuracy of the flux densities of the planets is only currently about 5-10% (with a few exceptions), so active research is required in this field.

ALMA will have the sensitivity to use much fainter sources as flux density standards than the current mm arrays. This opens the possibility to use new categories of objects. Among these, the brightest stars could be useful, especially at the highest frequencies. Asteroids or satellites of giant planets could also be interesting. For none of these objects is the prediction of the flux density accurate enough so far. ALMA should keep close contact with the communities of experts to improve this situation.

Objects like MWC 349, which has a well defined spectral index, can also be used provided an accurate measurement is made at the lowest frequency. Even quasars like 3C286, which have not shown any sign of time variations, could ultimately be used (possibly, though they begin to become more variable at mm wavelengths, since they become more core-dominated, and the cores are known to fluctuate – for example, 3C48 is a very good calibrator at frequencies less than about 15 GHz, but becomes quite variable by 40 GHz).. A detailed discussion of flux density standards is given in Yun et al. (1998) and Moreno & Guilloteau (2002).

Our current thinking is that we will have a very few true *primary standard calibrators*, whose flux densities are presumed known and stable – these may include Titan, Uranus, Mars (though it is rather too large to be most useful), the largest asteroids, and MWC 349, among others. We will then have a set of *secondary standard calibrators*, whose flux densities will have to be monitored relatively frequently (time scale of days) against the primary standards – these may include asteroids with diameter greater than about 200 km (so they are roughly spherical), and a number of stronger quasars. We will then have a large database of other possible calibrators, including quasars, stars, and other relatively strong, compact sources, which can be used to track the long timescale amplitude variations during each observation. These other calibrators must be measured against one of the primaries or secondaries each time they are used for an observation.

While more work is clearly needed in this area, it will be important for ALMA to keep a database of the calibration parameters, so that improvement in the calibration accuracy can be made when our understanding of the flux standards improves. ALMA should also keep close contact with other instruments working in similar frequency ranges which could provide key parameters on the flux density standards.

4.2.3 Phase Decorrelation

An uncorrected antenna based phase error of 10 degrees rms will result in a 3% decrease in the visibility amplitude due to decorrelation. As the characteristics of the phase noise change, the amount of decorrelation will also change. The primary defense against decorrelation is to try to correct for the phase (delay) fluctuations as much as possible. However, when the phase cannot be fully corrected, we can estimate the magnitude of the decorrelation and correct the visibilities. Decorrelation could be estimated from:

- phase calibrator data (fast switching)
- independent phase monitor (atmospheric) with an injected LO signal (antenna mechanical & electronic)
- radiometric data (atmospheric) with an injected LO signal (antenna mechanical & electronic)

Holdaway (1997c) has provided estimates of atmospheric coherence times. Electronic decorrelation is stable with time and (to first order) baseline length independent. Although it is antenna dependent, this should be calibrated out by relative amplitude calibration. Atmospheric decorrelation is the most damaging because of its baseline length dependence, but can be estimated from WVR data. Mechanical decorrelation remains to be evaluated. It is antenna dependent, but also possibly time dependent.

4.2.4 Changes in Transparency

At millimeter wavelengths, the changes in atmospheric transparency will be very modest, under 1% over 10 minutes about 80% of the time. Since the same amount of water vapor results in much larger opacities in the submillimeter, the transparency fluctuations in the submillimeter over characteristic calibration time scales will be much larger, typically several percent during median stability conditions. Because of the lack of strong enough amplitude calibration sources (Moreno & Guilloteau 2002), the transmission needs to be corrected either by appropriate prediction or measurement (see above discussion). Because of the many ozone lines in the submm windows, this transparency correction will need to be made with sufficient spectral resolution (10 to 100 MHz), and probably implies use of models for the highest spectral resolution data.

4.2.5 Polarization Complications

As mentioned below under polarization calibration, if a linearly polarized calibration source is used to track changes in the amplitude gain or opacity, a telescope with linear feeds will produce parallel hand visibilities which are modulated by the linear polarization. The extra signal varies as a sinusoid of the parallactic angle, so the errors are systematic. This must be accounted for.

4.2.6 Single Dish Issues

ALMA differs from any other aperture synthesis array in that, from the outset, the instrument will support no-compromise single-dish observing modes in addition to the more usual interferometric modes. Some of the issues are discussed in Emerson & Jewell (1993). Receiver stability and other issues are addressed in Welch (2003) and D'Addario (2003).

Because single-dish observing is in total power, albeit switched against, for example, blank sky, there are extraordinary demands on instrumental gain stability. In addition, the extra, variable emission from the sky comes in directly, and tends to mask the much weaker (by perhaps 4 orders of magnitude) astronomical emission. This is in contrast to interferometry, which by the use of cross-correlation rather than self-correlation, is relatively immune to these factors. Another serious problem for single-dish observing is that of the sideband gain ratio. Since some receivers will be double sideband (DSB), they will have emission from an unwanted part of the spectrum in the single-dish signal. The interferometry data do not suffer from this problem because the unwanted sideband is rejected in the downconversion.

Astronomical calibration in single-dish mode has to be on a dish-by-dish basis; calibration sources need to be detectable with adequate signal-to-noise ratio by one single dish of the array. This is again in contrast to interferometric astronomical calibration measurements, in which the large collecting area of the entire array can contribute to the signal-to-noise ratio achievable in calibrating individual dishes of the array. We can attempt to calibrate interferometrically, and apply these gains to the single-dish data, but in this case, the decorrelation must be accurately calibrated, as must the sideband gain ratio.

Polarization calibration of single-dish observations has its own problems. At mm wavelengths, polarization measurements are conventionally made with a "widget" in front of the receiver feed. This "widget" introduces changes in the polarization response of the receiver - for example a rotating grid and screen combination can continuously rotate the incident plane of linear polarization. The astronomical polarization is then detected by synchronous changes in total power intensity through the receiver as the sense of polarization changes.

ALMA may indeed have to provide such "widgets" for each of the antennas. However, the complexity and potential unreliability of such a device could be avoided if it were shown possible to measure polarization reliably, in single-dish mode, by cross-correlation of the signal from orthogonally polarized feeds. Tests of the feasibility of this techniques are planned.

4.2.6.1 Atmospheric Emission Cancellation

The emission from the atmosphere is much stronger than the emission from most astronomical sources, and, even worse, the atmospheric emission is variable as well. The variable part of the emission is mainly due to inhomogeneously distributed water vapor, which also causes the phase fluctuations. Since we have excellent statistics of the phase stability on the Chajnantor site, we can infer the severity of the variable atmospheric emission at any desired frequency by using a transmission model or FTS measurements.

For an interferometer, the atmospheric emission above two different antennas is not correlated, so it does not affect the visibilities. In total power continuum observations, the variable atmospheric emission is a major problem which requires some sort of switching on the sky. The total power spectral line case is much less demanding, as large atmospheric fluctuations can be tolerated, considering the much smaller channel widths and much higher thermal noise and the possibility of fitting an average baseline to each spectrum. The spectral line data will have secondary effects, such as the bandpass changing in response to the changing atmospheric load. However, the spectral line observations are much easier than the continuum case, so if we can beat the atmosphere for continuum observations, the spectral line observations will be no problem. The detailed treatment of this problem is presented in an upcoming ALMA Memo (Holdaway, Lugten, and Freund, in preparation). Lucas (2000) presents a novel method for calibration to remove atmospheric emission which makes use of the very wide bandwidths of ALMA.

4.2.6.1.1 Beam Switching

Traditionally, beam switching by a nutating subreflector has been used to remove the variable atmospheric emission. Our study indicates that most beam switching is non-optimal. For any given observation, we would like to be roughly equally limited by thermal SNR and by the residual variable atmospheric emission. If the noise is dominated by the variable atmospheric emission, we need to switch faster. The faster we switch, the better the atmospheric cancellation, but the lower the duty cycle, so the thermal noise will increase. Furthermore, the distance of the throw also needs to be considered. In general, it is optimal to have the smallest throw which gets completely off source. However, in an unstable atmosphere, multiple short throws are better. Hence, the detailed use of a nutating subreflector needs to be fine tuned to match the atmospheric conditions and the observing frequency. As with fast switching, we hope that the observer does not have to perform the calculations to find the optimal switching strategy; the observer should provide high level guidelines, and the program which performs the micro-scheduling should calculate the optimal switching strategy for the current atmospheric conditions.

The nutator design (see specifications in paragraph 4.2.9 of the Antenna chapter of this ALMA Construction Project Book) for the ALMA prototype antenna allows for a maximum throw of about 1.5 arcmin for symmetric beam throwing. Maximum nutating frequencies of about 10 Hz are planned. If it is affordable, nutators with higher peak acceleration and larger maximum throws would be desirable for the production antennas. The two beams should be as similar as possible, and as similar as possible to the “non-wobbled” beam, to reduce the level of systematic errors in beam switching. All of these beams must be accurately measured and recorded.

The analysis of the On-The-Fly technique for total power continuum observations indicates that it will be as good or better than beam switching in all situations. However, there is considerable risk involved in relying on the On-The-Fly method to cancel all atmospheric fluctuations. For this reason, it is generally agreed that the prototype antennas need to have nutating subreflectors. Currently, this is planned according to the Payne design referred to above.

4.2.6.1.2 On-The-Fly

In On-The-Fly (OTF) observing, the antennas scan quickly across a source at constant elevation angle, using the off-source regions on other side of the source region to define the sky emission. Very large sources will need to be pieced together at some SNR expense. The OTF technique promises to be quite effective at removing the atmospheric emission for three reasons:

- each Nyquist sample on the sky is observed for a very short time, so the system noise is large and a larger amount of sky fluctuation noise is tolerable. (The large number of Nyquist samples observed in each scan compensates for this large noise per Nyquist sample.)
- since more time is spent observing the OFF than an individual ON Nyquist sample, the atmosphere is well determined, unlike beam switching where we are differencing two noisy numbers.
- since the OFF's are observed over a range of time, we can remove a second order polynomial trend in the atmospheric emission time series, which greatly reduces the residual sky emission fluctuations.

For sources which are about one beam across, the OTF observing strategy works about as well as beam switching. For larger sources, OTF wins because of the relative increase in the SNR of the atmospheric determination and because multiple throws begin to degrade the beam switching SNR.

Because the entire antenna is moving, many systematic errors which plague beam switching (such as differences in the shapes and gains of the ON and OFF beams) are eliminated. However, it takes much more energy to move the entire antenna, and there is more risk in general with an observing strategy that attempts to move the entire antenna.

OTF will work only if we can slew and reaccelerate the antenna quickly without exciting the lowest resonant frequency of the antennas. An initial analysis of this problem has been performed by Holdaway, Lugten, and Freund (in preparation). Using a Gaussian acceleration profile and an error function velocity profile, they predict the antennas will be able to turn around from one scan direction to the other in about 0.2 s without appreciably exciting the lowest resonant frequency. This acceleration profile is a good one, but probably not an optimal one, so further work could help optimize the profiles for both OTF antenna motion and fast switching antenna motion. In order not to excite the antenna motions, the acceleration must be very smoothly varying. This will put strong constraints on both the control system and on the servo system.

OTF simulations of sources of various sizes indicate that the optimal slew velocity varies linearly with source size. For a maximum interesting source size of 1 deg, a maximum slew rate of about 0.5 deg/s is required. This requires a maximum antenna angular acceleration of about 12 deg/s/s. Since the profile is Gaussian, we do not require this maximum acceleration for very long. These maximum velocities and accelerations are for an antenna with lowest resonant frequency of 6

Hz. An antenna which was less stiff could not utilize such large accelerations and velocities in OTF observing. A stiffer antenna would permit faster turnarounds, requiring larger accelerations and velocities. However, the 6 Hz antenna is effectively beating the atmosphere already, so not much is gained from a stiffer antenna.

OTF requires that we know where the antenna is for each Nyquist beam. At the 0.5 deg/s maximum slew rate, observing at 850 GHz with a half beam size of 0.001 deg will require that we dump the data and know where the antenna is every 2 ms. We don't need to make the antenna go to any precise place at any precise time, we just need to know where the antenna was at a precise time. We may not need to read the encoders every 2 ms; if the antenna position changes smoothly over time scales of 10 ms, we can read out the encoders more coarsely and interpolate. We do not require that the encoders be accurate to within the pointing specification of 0.6 arcsec.

For total power OTF, we need not synchronize all the antennas in their dance across the sky. The antennas could be staggered to permit a more constant utilization of electrical power.

4.2.6.1.2 1/f Noise

In addition to atmospheric brightness fluctuations, beam switching and OTF will remove a portion of the receiver 1/f noise. From the optimizations we have performed, we can set specifications on the 1/f noise for each observing frequency. Even though the beam switching is performing the switching faster than OTF, the integration time spent on each ON is often larger than the integration time spent per Nyquist sample of an OTF observation, so OTF and beam switching are similar in their ability to switch out 1/f noise. If these specifications cannot be met, we must reoptimize the OTF observing strategy, which would result in moving more quickly to accomplish faster switching and less time or more white noise per Nyquist sample on the source. This would favor both higher maximum accelerations and a stiffer antenna. D'Addario (2003) has discussed some of the details of 1/f noise, and compared OTF with beam-switched measurements.

Freq	Beam Size Source		0.5 deg Source	
[GHz]	noise [Jy]	break frequency [Hz]	noise [Jy]	break frequency [Hz]
90	0.047	1.2	0.081	0.34
230	0.088	1.2	0.25	0.29
345	0.14	1.2	0.47	0.29
650	0.33	1.3	1.6	0.34

Table 1: For continuum (8 GHz bandwidth per polarization) OTF observations, what noise level must the 1/f noise be below, and at what frequency, for 1/f noise to have essentially no effect on OTF observations' sensitivity?

4.2.6.2 Sideband Gain Ratio

An important issue in the single-dish calibration is the fact that for DSB receivers, there is emission from both sidebands which is detected. This is not an issue for the cross correlated visibility data, because of the phase switching in the LO which does effective sideband rejection (D'Addario 2001). Moreno & Guilloteau (2002) have studied this in a preliminary way, and find that this may be one of the limiting calibrations in terms of the time it takes to properly make the measurement. During spectral line observations (especially surveys), it may be possible to deconvolve the ratio, but in normal continuum observations, this will be nearly impossible. This needs more study, and some experience on the telescopes.

4.3 Delay

4.3.1 Atmospheric Delay

4.3.1.1 Neutral Atmosphere

Most of the systematic and long-term fluctuating part of the atmospheric delay is taken out by observing nearby calibrators on timescales of minutes or even slightly longer. We assume that this will be true for ALMA, and so have no special scheme for calibrating them out. In addition, if there are fast fluctuations in the dry component of the atmosphere, we assume that the fast switching will account for them (see below). This, however, needs some more attention – both as a theoretical study, and with experience on the telescopes.

The main part of the calibration of the fluctuations in the neutral atmosphere is dealing with fluctuations in tropospheric water vapor. We currently have two planned techniques for calibrating these fluctuations, fast switching, and water vapor radiometry.

4.3.1.1.1 Fast Switching

If a calibrator is sufficiently close and the telescope is sufficiently fast, fast switching between a calibrator source and a target source can effectively stop the atmospheric, electronic, and antenna phase fluctuations. If fast switching is used as the phase calibration method, it makes minimum requirements on the system sensitivity, the slew speed and settle down time of the antennas, and the online and data taking systems. Fast switching has been studied extensively in the MMA and ALMA memo series (see especially Holdaway 2001b; Holdaway & Pardo 2001; and references therein), and we are fairly confident that it will work for ALMA.

The basic criteria for fast switching to work is that the phase calibration source needs to be detected with sufficient SNR and the target source be observed for some amount of time within the coherence time and distance of the atmosphere. This translates into a requirement that there be sufficiently many calibrator sources which are sufficiently bright (Holdaway, Owen, & Rupen 1994; Foster 1994), and a requirement on the sensitivity of the array. In practice, this means that the calibrator source will typically be within a few degrees of the target source, the calibrator will usually be detected in less than a second, and the entire cycle time will be about 10 s, though the details vary with observing frequency. Spectral line observations will need to use wide bandwidth continuum observations of the calibrator.

With the current sensitivity of ALMA and our understanding of the quasar source counts and their dependence on frequency, we will not always be able to perform fast switching calibration at the target frequency, but often we will get a higher SNR phase solution by observing the calibrator at a low frequency (90 GHz) and scaling the solution up to the target frequency. This is not to imply that fast switching will *always* use a lower frequency for the calibrator, but that sometimes it will prove to be more effective. The optimal frequency for calibration has been computed by Moreno & Guilloteau (2002). If the calibrator is observed at a lower frequency, and the phases scaled up, then this requires a much more accurate phase solution at the lower frequency. Since the dry atmosphere is non-dispersive, this extrapolation basically relies upon the wet differential delay to be non-dispersive as well. In the submillimeter, the wet differential delay is dispersive, which will either limit the effectiveness of fast switching or require more complications in the fast switching observing strategy, such as less frequent multi-frequency calibrator observations to help separate out the non-water vapor phase contributions. These multi-frequency observations may also be required for amplitude calibration. On longer timescales, set by the phase stability of the electronics and mechanics, it is necessary to measure the instrumental phase and delay offsets between different frequencies, by observing a single source at both wavebands.

The antenna movement requirement from fast switching is currently a slew of 1.5 degrees and settle down to 3 arcsec pointing in 1.5 seconds.

The on-line system needs to control the antennas gracefully enough to move them quickly without exciting the lowest resonant frequency. Also, the quanta of integration time and scan length need to be sufficiently small so as not to restrict the integration time spent on the target source and calibrator or the time spent between sources. Flexibility at the 100-200 ms level is desirable. Fast switching data can be calibrated with existing software, but some extensions in spatial-temporal interpolation will be useful.

Fast switching will reduce the sensitivity of observations due to time lost observing the calibrator and moving the antennas, and due to decorrelation from residual phase errors. Both effects can be reduced by observing in the best conditions, which often result in very low residual phase errors at a minimum expense in time lost to the calibration process . However, not all projects can be observed during the best phase conditions. Holdaway (2001c) concludes that fast switching will generally result in less than a 20% decrease in sensitivity for the phase conditions at the Chajnantor site.

During poor phase stability conditions, fast switching won't work at the high frequencies. Also, a given target field may have a dearth of calibrator sources, requiring that the field be observed during better phase conditions than the average field. For reasons like these, dynamical scheduling is absolutely required to optimize the utility of ALMA. We envision one or more phase stability monitors providing real time information to the array control center, and contributing to observing decisions - e.g.:

- what project should run on the telescope?
- do the present conditions permit the current project to continue?
- what is the optimal calibrator for the current project in the current atmospheric conditions and hour angle?

The quasars which will form the bulk of the fast switching calibrators will be highly variable at millimeter wavelengths, and a quick survey of a few square degree region about the target source will sometimes be required. ALMA has the sensitivity to perform a blind search for calibration sources in a few minutes. Surveys directed with lower frequency source catalogs will be even faster. Whenever a potential calibrator is observed, the source information will need to go into a comprehensive calibrator database, which can also be used for choosing an appropriate calibrator.

4.3.1.1.2 Water Vapor Radiometry

The most promising alternative to fast switching is radiometric phase correction or, more specifically, Water Vapor Radiometry (WVR), which has a thorough treatment in the MMA and ALMA memo series (see especially Hills & Richer 2000; Hills et al 2001; and references therein). WVR utilizes the variable emission caused by inhomogeneously distributed atmospheric water vapor to determine the phase fluctuations caused by water vapor. While water vapor is not the only source of phase errors, it is the dominant source of short time scale phase fluctuations. This method has had several early successes, but the correlation between the radiometric fluctuations and the interferometrically measured phase fluctuations changes with time, and there are some times when the method does not work well at all.

The baseline plan for radiometric phase correction is a 183 GHz water vapor radiometer using an uncooled Schottky receiver with 5 (or 2 times 5) frequency channels. The radiometer may be of the Double Dicke-switch type or of the correlation type, depending on the results of the development phase. The optics system is designed to allow using the radiometer simultaneously with all observing bands, but not during some calibration operations (notably when the amplitude calibration widget is in use).

Details of the WVR development strategies are developed in Hills & Richer (2000) and Hills et al. (2001). Active research in the interpretation of the WVR measurements is mandatory. Before WVR prototypes are available for on-sky tests, preferably at the ATF, but possibly on other interferometers (e.g. IRAM or SMA), extensive simulations should be performed using the ATM model. The ALMA simulator developed by the ADACE consortium incorporates such simulation capabilities: early delivery of the development versions of this simulator to the Science IPT would be valuable.

4.3.1.1.3 Paired Antenna Calibration

It is possible to use some of the antennas to observe a calibrator and the rest of the antennas to observe the target source. In the smaller arrays, the configurations will naturally permit paired array calibration. Carilli & Holdaway (1999) present encouraging results using the VLA. Longer configurations have no specific provision to optimize use of such a mode. This mode should be reserved as a backup solution for special projects in which the resulting sensitivity loss is not a major issue, but will not be the default for ALMA.

4.3.1.2 Ionosphere

It may be necessary to track fluctuations in the ionospheric delay (see discussion above). If this is the case, then it is likely that we will need some sort of GPS system which can track the total electron content, and possibly in combination with some ionospheric model, can estimate ionospheric delay as a function of pointing location (see, e.g., Gradinarsky et al. 2001).

4.3.2 Calibration of the Electronic and Antenna Structure Delay with an Injected Signal

Radiometric phase correction will only correct for those phase fluctuations which are caused by water vapor, and will not correct for any phase errors caused by variations in the dry atmospheric delay, mechanical instabilities in the antenna, or instabilities in the electronics. Therefore, radiometric phase correction requires some supporting observations or calibration technique to remove phase errors caused by these other sources.

An initial idea was to periodically inject a stable signal, perhaps derived from the LO, into the feed to calibrate the electronic contributions to the phase errors. If the calibration signal is injected from the subreflector, then this calibration system will also track the mechanical phase drifts due to focus and surface changes of the antenna. If the calibration signal is derived from the LO, and the LO itself has phase instabilities, they will either cancel or be doubled, depending upon the relative phase of the LO and the injected signal. In fact, by alternating the relative parity of the injected signal and the LO, we can solve for both phase errors in the LO and in the rest of the electronics and the antenna up to the subreflector. However, phase changes due to mechanical deformations in the antenna mount (i.e. baseline fluctuations, for example due to wind-induced tilt of the fork) are not accounted for in such a system. Fluctuations in the dry atmosphere will also be unaccounted for.

The on-line system would need to control the details of the injected signal. Information about the injected signal would need to be recorded with the data, and an option for determining and correcting for the electronic phase errors in real time should exist.

The injected signal calibration scheme is still an area of research in this phase of the project. Implementation decision can only be taken once a prototype is build and tested at the ATF.

4.4 Bandpass

When observing spectral lines, we would like to always be in the situation where the noise in the channel is limited by thermal noise rather than bandpass calibration. A majority of the spectral line observations made with ALMA will probably have no problem meeting this condition.

In rough terms, the dynamic range DR of a single spectral channel which is limited by errors in continuum subtraction caused by bandpass errors will be:

$$\mathrm{DR} = \frac{S_{\mathrm{line}}\sqrt{N_{\mathrm{ant}}}}{S_{\mathrm{cont}}\Delta s} \tag{1}$$

where S_{line} is the flux density of the peak of the line, N_{ant} is the number of antennas, S_{cont} is the continuum flux density of the source, and Δs is the rms error in the bandpass. For the spectral line observations to be limited by thermal noise and not by bandpass errors, and assuming the bandpass errors are themselves due to thermal noise in the observations of the bandpass calibrator, we have the condition that

$$N_{ant}\sqrt{\frac{t_{cal}}{t_{line}}} > \frac{S_{cont}}{S_{cal}}$$

where S_{cal} is the flux density of the bandpass calibrator, t_{cal} is the time spent on the bandpass calibrator, and t_{line} is the time spent on the line source.

Because in some cases, the continuum flux density of the line source is quite large (planets and strong quasars, for example – see discussion in requirements section), we cannot simply rely on the technique of choosing a bandpass calibrator which is much stronger than our line source. Injection of a strong coherent (so that the interferometric mode can be used) noise signal which is flat over the observed frequency range would be an ideal solution to both the planet problem and to the galactic confusing line problem. However, such a system has not yet been successfully developed so far, and the requirements (flatness of frequency response and time dependence) are extremely difficult to meet.

We still always have the capability of calibrating the bandpass by observing a strong source for a sufficiently long time. Because the noise increases with frequency, while the calibrator source strength usually decreases, it will be advantageous for SNR considerations to split the bandpass calibration into several separate pieces: wideband at the observing frequency (WO), wideband at the calibration frequency (WC), and narrowband at the calibration frequency (NC). The final bandpass (narrowband at the observing frequency, or NO) is then given by:
NO = NC x WO / WC.

At millimeter and submillimeter wavelengths, the atmosphere will also contribute to the bandpass for wide bandwidth observations, so we must either perform an independent determination of the bandpass astronomically to solve for the atmospheric bandpass component, or we would measure the precipitable water vapor from opacity measurements made at a fiducial frequency and determine the atmospheric contribution to the bandpass through the use of an atmospheric transmission model. Currently, the atmospheric transmission models are probably not good enough for this sort of work, but ALMA would provide enough data for an ad hoc model or to improve the theoretical models. We will be concerned with changes in the atmospheric component of the bandpass on reasonably short time scales and among the different antennas. Prediction of the magnitude of such changes is feasible using the ATM model. It is essential to note that because of a large number of Ozone lines in the sub-mm domain, the bandpass requires modeling at quite high resolution, some of the lines having large opacity changes on scales of a few MHz only.

For pure continuum observations, masking these regions of the frequency spectrum may be needed in software to improve the amplitude calibration. In some amplitude calibration techniques, the calibrated visibility is derived from the measured correlation coefficient and the calibrated auto-correlation spectra (i.e., Tsys as function of frequency). For Single-Sideband systems, this should provide an accurate bandpass calibration. The situation for Double-Sideband systems is more complex because of the emission from both bands, and should be investigated in details.

There is an implicit specification placed on the system design that the electronic bandpass be either stable or that it vary linearly with time to something like 10000:1 to 100000:1. If the bandpass changes are mainly linear, we can remove them through interpolation if we calibrate often enough.

A significant difficulty in the stability of the bandpass is the existence of standing waves in the system. Standing waves may be in emission (which affect bandpass normalization) or in transmission (which directly affect the astronomical signal). Any modification of the optical layout of the antenna will modify the standing wave pattern. This will happen in the dual-load calibration system when switching from load to signal for example. Standing waves will usually average out as $\sqrt{N_{ant}}$ since every antenna has a different standing wave pattern. An issue which remains to be quantified is what design of the calibration system will minimize the standing wave problems.

4.5 Polarization

ALMA will use linearly polarized feeds because they have a wider usable bandwidth than circularly polarized feeds, and can provide complete coverage of all millimeter wavelength atmospheric windows with a reasonable number of receivers. Cotton (1998) treated the problem of polarization calibration for the MMA in detail. A more recent treatment for ALMA can be found in Appendix C to the Report of the ALMA Science Advisory Committee March 2000 Meeting. The general problem of calibrating polarization in interferometry has been described fully by Hamaker, Bregman, & Sault (1998), who develop a Jones matrix approach (for an earlier similar approach, see Schwab 1979). The main detail that we must be concerned with here is that the measurement of linear polarization is corrupted by contamination from Stokes I. For linearly polarized feeds, this corruption is in the form of a gain stability term (as opposed to circularly polarized feeds, where the corruption arises from a leakage term). Another point of note is that it is not easy to distinguish circular polarization from the instrumental polarization terms when using linearly polarized feeds.

Because the linear polarization is entangled with the total intensity, there are times when all four cross correlations per baseline will need to be performed, which will probably result in halving the bandwidth and cutting the sensitivity by root two. We consider several cases which could come up with ALMA to demonstrate when we may need to consider all four cross correlations and when we may use approximations to make use of just the two parallel hand cross correlations:

- Amplitude calibration is performed by knowing precisely the gains and system temperatures of the antennas (not by looking at an astronomical source), and phase (delay) calibration is performed on a quasar (or a combination of radiometric [WVR] plus a quasar). The quasars will generally be a few percent linearly polarized, but may be as much as 10-20% polarized, and hence Stokes Q and U will influence the parallel hand visibilities. These sources have almost no circular polarization. For a point source, the linear polarization of the calibrator will not affect the phase, only the amplitude. We further consider two subcases:

  - Total intensity imaging with no polarization in the target source. Many millimeter spectral line sources will have little or no linear polarization. Nothing special needs to take place, as the parallel hands will basically contain Stokes I.
  - Total intensity imaging with appreciable linear polarization in the target source. The linear polarization in the target source will corrupt the parallel hand visibilities in a systematic way. However, when the XX and YY visibilities are added together, the linear polarization corruptions cancel out. This is acceptable for low to moderate dynamic range total intensity observations, but may not be sufficient for high dynamic range total intensity observations, as residual gain errors will limit the cancellation of the linear polarization and adding the XX and YY correlations results in a condition in which gain errors no longer close, limiting the use of self-calibration. High dynamic range total intensity imaging of a source with appreciable linear polarization may require full polarization calibration and imaging.
  - Polarization imaging. A bright calibration source must be observed to determine the instrumental polarization leakage or "D" terms. If the calibrator has known (or zero) linear polarization and no circular polarization, the D terms can be determined in a single snapshot. If the calibrator has unknown linear polarization, the calibrator must be observed through sufficient parallactic angle coverage to permit separation of the calibrator and the D terms. Application of the D terms will permit the polarization imaging.
- Amplitude calibration is performed by assuming the flux density of some astronomical sources is known precisely, and using measurements of that source to set the voltage to flux density conversion scaling. If the source of flux density is not polarized, there is no problem. If it is linearly polarized, then the parallel hand visibilities will vary systematically with parallactic angle, the XX and YY visibilities varying in opposite senses. There are several options:
  - For total intensity observations of a target source at low to moderate SNR, the array-wide XX and YY gain ratios can be determined and corrected for.
  - High SNR total intensity observations will require accounting for the different parallactic angles of each antenna, which will result in imperfect cancellation when using the array-wide gain ratios. In this case, the full polarization calibration will need to be performed on the quasar, even if there is no interest in polarization. Full polarization observations of the source are only needed if desired.

In all cases in which the cross hand visibilities are explicitly used, the X-Y phase offset must be monitored for each antenna. As there is no simple way to determine the X-Y phase offset astronomically, ALMA could inject a tone into the feeds, as the AT does. Cotton (1998) points out that it is difficult to generate a millimeter RF tone, and that injecting an IF tone further downstream in the electronics is simpler, though not as good instrumentally (it does not calibrate the portion of the offset which occurs before the IF). On the other hand, we could derive an RF signal from the LO and inject it into the feeds for the X-Y phase calibration.

The choice of a flux density calibrator may also interact with the polarization calibration. Unresolved asteroids which are not azimuthally symmetric will have some time dependent linear polarization, which will complicate the flux density calibration. If stars are used for a flux standard, they may display some circular polarization, which would require that another source be used for the D term calibration.

As stated above, the full polarization calibration requires good coverage in parallactic angle to separate the constant instrumental polarization (D term) signal from the sinusoidally varying astronomical polarization signal. This causes some concern since ALMA is envisioned to be predominantly a near-transit instrument with real time imaging capability. If instrumental polarization calibration is required for many observations, it may be prudent to keep a database of the instrumental polarization solutions at the various frequencies and bandwidths and rely upon that whenever possible. Unlike the VLA, the ATNF compact array shows essentially no time variability in the instrumental polarization (less than 1:10000 over 12 hours, with variations of 0.1% over months). Given the constraints of ALMA, time constant instrumental polarization is certainly an important design goal. An analysis of the cause of the VLA unstability would also be useful.

One way around the complication of good parallactic angle coverage is to use sources of known polarization (one special case of which is totally unpolarized sources). Holdaway, Carilli, and Owen (1992) have demonstrated that it is possible to solve for the instrumental polarization for a single snapshot, (i.e., a single parallactic angle) if the source polarization is known in advance. So, it would be beneficial to ALMA observing to identify bright, compact sources with known polarization or no polarization for use as polarization calibrators. Unfortunately, such sources are currently completely unknown at millimeter wavelengths – all quasars have variable polarization angle. Some study of what level of polarization is acceptable as “unpolarized” is warranted.

6 Miscellaneous

6.1 Solar Observing

Calibration of solar observations will involve special care. When observing the sun, a special attenuator will be placed in front of the receiver package. Because of this, it will be necessary to calibrate the properties of that attenuator. Normal calibrator-source-calibrator sequences will be of little use, because there are few calibrator sources which are strong enough to be detected through the attenuator, and if they are observed without the attenuator in place, there will be an unknown amplitude and phase effect from the attenuator which must be calibrated.

References

Thompson, Moran, & Swenson, Interferometry and Synthesis in Radio Astronomy, 2nd Ed., Wiley, 2001
Fomalont & Perley, in SIRA II
Cornwell, T.J., M.A. Holdaway, & J.M. Uson, Radio-interferometric imaging of very large objects: implications for array design, A&A, 271, 697, 1993
Holdaway 1997a, Memo 178
Morita 2001, Memo 374
Pety et al. 2001a, Memo 387
Pety et al. 2001b, Memo 398

Yun et al. 1998, Memo 211
Moreno & Guilloteau 2002, Memo 372
Woody et al. 1995, Memo 144
Hills & Richer 2000, Memo 303
Sramek 1981 (VLA Scientific Memo 133)
Sovers, O.J., J.L. Fanselow, & C.S. Jacobs, 1998, Rev. Mod. Phys., 70, 1393
Clark 1973 (VLA Computing Memo 105)
Butler 2003 (unpublished note from 1998-Dec-10 on axis intersection for MMA)
Wade 1974 (VLA Test Memo 104)
Holdaway 2001 (Memo 402)
Wright 2002 (Memo 427)
D'Addario 2000 (Memo 298)
Holdaway 1992 (Memo 74)
Lamb 1993 (Memo 96)
Ruze, J., 1969, "Small Displacements in Parabolic Reflectors", unpublished
Lamb & Forster 1993 (Memo 100)
Lamb & Woody 1998a (Memo 234)
Woody & Lamb 1999 (Memo 241)
Plathner et al. 1999 (Memo 254)
Anderson 1999 (Memo 253)
Kingsley et al. 1999 (Memo 263)
Plathner 1999 (Memo 259)
Lugten et al. 1999 (Memo 240)
Lugten 1998 (Memo 232)
Payne 1997 (Memo 181)
Mangum 2001 (Memo 366)
Bayley, A.M., R.J. Davis, J.S. Haggis, & H. Kärcher, 1994, A&A, 283, 1051
Greve, A., J.-F. Panis, & C. Thum, 1996, A&AS, 115, 379
Scott, S., 1993, "Radio Pointing Algorithm", OVRO Millimeter Array Memo
Olmi, L., & M.M. Davis, 1998, A&AS, 129, 177
Lucas et al. 2003 (Computing Memo 11)
Mangum 2000 (Memo 288)
Wright 1997 (Memo 180)
Holdaway 1993 (Memo 95)
Holdaway 1996 (Memo 159)
Lucas 1997 (Memo 189)
Rupen 1997 (VLA Test Memo 202)
Sakamoto 2000 (Memo 322)
Holdaway 1997b (Memo 186)
Butler 1997 (Memo 188)
Holdaway and Woody 1998 (Memo 223)
Lamb and Woody 1998b (Memo 224)
Mangum 2000 (memo on optical pointing telescope)
Snyder et al. 2000 (Memo 315)
Otárola et al. 2002 (Memo 413)
Jacobs & Rius 1990, JPL IOM 335.6-90-034
Ryan, J.W., & Ma, C., in Radio Interferometry Techniques for Geodesy (NASA Conf. Publ. 2115), 347, 1979
Butler, B.J., "Requirements for Subreflector and Feed Positioning for ALMA Antennas", unpublished note, 2003-May-08
Otárola, A., & M.A. Holdaway, 2002, Memo 419
Wade 1976 (VLA Scientific Memo 122)
D'Addario, L., R. Bradley, E. Bryerton, R. Sramek, W. Shillue, and S. Thacker, ALMA Project Book, Chapter 7, Local Oscillators, 2002
Hales et al. 2003, Memo 459
Scott, P.F., & M. Ryle, A rapid method for measuring the figure of a radio telescope reflector, MNRAS, 178, 539, 1977

Johnston, K.J., & Chr. de Vegt 1999, ARA&A, 37, 97
Veissel-Vernier, M. 2003, A&A, 403, 105
Bock et al. 1998, Memo 225
Mangum 2002, Memo 434
Emerson & Jewell 1993, Memo 108
Welch 2003 – memo 454
D'Addario 2003 – memo 466
Hills et al. 2001, ALMA Memo 352
Houde, M., R.L. Akeson, J.E. Carlstrom, J.W. Lamb, D.A. Schleuning, & D.P. Woody, Polarizing Grids, Their Assemblies, and Beams of Radiation, PASP, 113, 622-638, 2001
Cotton 1998, MMA memo 208
Hamaker, Bregman, & Sault 1998
Schwab 1979 – VLA computing memo 154
Sault et al 1991
Holdaway, Carilli, and Owen (1992, VLA Scientific Memo 163)
Bennet & Godwin 1977
Butler 1999, VLBA Test Memo 62
Ulich & Haas 1976
Kutner 1978
Guilloteau 2002, Memo 422
Bock, D., & J. Welch 2003, unpublished memo "Measurements of the prototype dual-load calibration system on BIMA"
Martin-Pintado, J., S. Navarro, S. Martin, & M. Carter 2003, unpublished memo "Results of the tests of the semitransparent vane"
Houde, M., R.L. Akeson, J.E. Carlstrom, J.W. Lamb, D.A. Schleuning, & D.P. Woody 2001, Polarizing Grids, Their Assemblies, and Beams of Radiation, PASP, 113, 622
Holdaway 1997c, Memo 169
Lucas 2000, Memo 300
D'Addario 2001, Memo 385
Holdaway 2001b, Memo 403
Holdaway & Pardo 2001, Memo 404
Holdaway, Owen, & Rupen 1994, Memo 123
Foster 1994, Memo 124
Carilli & Holdaway, Memo 262
Gradinarsky, L.P., J.M. Johansson, G. Elgered, & P. Jarlemark 2001, GPS site testing at Chajnantor in Chile, Phys Chem Earth A, 26, 421
Guilloteau, S., & A. Bacmann, Memo 461
Pardo, J.R., J. Cernicharo, & E. Serabyn, Atmospheric Transmission at Microwaves (ATM): An Improved Model for mm/submm Applications, IEEE Trans. Ant. Prop., 49, 1683-1694, 2001