Very Long Baseline Interferometry and the VLBA
ASP Conference Series, Vol. 82, 1995
J. A. Zensus, P. J. Diamond, and P. J. Napier (eds.)

Printed June 13, 1995

CHAPTER 19

Astrometry

EDWARD FOMALONT

NATIONAL RADIO ASTRONOMY OBSERVATORY

Abstract The special VLBI techniques used to obtain accurate absolute or relative positions of radio sources are discussed.

Contents

19.1 Types of Astrometry

The major goal of radio arrays like the VLBA and the VLA is the formation of an image in a region of sky. Two questions are generally asked about any object or objects (called radio sources) in the image: (1) What does the radio source look like? and (2) What is the location of the radio source? The first question deals with the shape of the radio source with the goal of the observations, calibrations and imaging to produce a high quality representation of the radio source, with as much instrumental and environmental distortion removed as possible. This is called *imaging.* The precise location of the radio source is, often, secondary.

Many astrophysical questions deal specifically with the location and motion of celestial objects, rather than their shapes, and this area of astronomy is called *astrometry.* Astrometric observations can be divided into two parts: **absolute** astrometry attempts to determine the position of an object relative to *a fundamental reference frame;* **relative** astrometry determines the relative positions among several objects over a small region of sky, or the internal motions within a radio source. Because astrometric observations often deal with small diameter radio sources, images need not be produced, because the positional information can be obtained from direct analysis of the (u, v) data. However, to determine positional information amongst extended sources, both imaging and astrometric techniques must be used.

19.1.1 Absolute Astrometry

Absolute astrometry is associated with fundamental astrometric and geodetic definitions because the measurement of absolute celestial coordinates must deal with inertial reference frames, earth rotation, and orientation, the deformation and distortion of the surface of the earth and the telescope, gravitational interaction of solar system bodies, especially with the moon (tides), and propagation characteristics of the atmosphere and ionosphere. Often, the absolute astrometric problem is turned around; for example, the use of VLBI for geodesy.

By using radio sources as beacons in the sky, information about the geodetic framework needed to define the locations of the telescopes in the experiment can be obtained. In fact, most accurate astrometric projects require simultaneous solution of absolute positions with other important earth-related parameters.

19.1.2 Relative Astrometry

A sufficiently accurate position of a radio source can generally be obtained by measuring its position with respect to convenient nearby objects (calibrators), typically a few degrees away. The accuracy of the derived position then depends on both the quality of the calibrator positions and the tie-in error between the source and calibrator(s). Relative positions depend much less on the uncertainties associated with absolute astrometry because their effects are reduced over small regions of sky. However, relative astrometry at the micro-arcsecond level still requires competent absolute astrometric calibrations. A natural extension of relative positions is the measurement of the kinetic properties of an object: proper motion, parallax, and orbital trajectory in the case of a multiple system or a solar system object. If the positions are tied to the same calibrator(s) for each observational epoch, then the accuracy of the motion is nearly independent of the quality of the calibrator position, as long as the calibrator is stationary.

The measurement of internal motions in complex radio sources—superluminal motions in jets of quasars, shocks or blobs in radio galaxy jets and lobes, storms in planetary atmospheres—is more an imaging problem than an astrometric problem because high fidelity images must be made in order to detect the changes within the source. Some interesting image and visibility data techniques have been developed to determine the motions within complex sources (Fomalont et al. 1992; Biretta and Owen 1990).

Because all radio sources are extended and many are variable in intensity and shape, absolute and relative astrometric accuracy is often limited by the internal radio structure. With sufficiently detailed observations, radio source shapes can first be obtained using normal imaging techniques. Then, the astrometric analysis can incorporate the effects of the finite size of the radio source.

19.2 The Basic Array Measurement

19.2.1 The Total Phase Delay

The normal response of a two-element interferometer, $r(t)$, has a quasi-sinusoidal time behavior, called *fringes,* with the form

$$r(t) = A \cos \phi_T \,, \tag{19.1}$$

where A is the amplitude of the fringe pattern and ϕ_T is the *total phase delay* between the radio signal impinging on the two antennas. The phase delay can be many thousands of radians in magnitude and A and ϕ_T are functions of time and direction. Using the myriad of calibration and supporting measurements, they contain all of the information needed to image a radio source and to determine its precise location in the sky.

The amplitude depends on the receiver and antenna gain characteristics and the propagation attenuation, G_0 and the bandwidth of the received signal $\Delta\nu$. The intensity of the radio emission in the field of view of the antennas, including the effects of the shape of this emission, is called the *visibility amplitude* of the source, V:

$$A = G_0 \Delta\nu V\,. \tag{19.2}$$

All of the fringe amplitude components are functions of time and direction. In most astrometric applications, the fringe amplitude is of secondary importance.

The major component , ϕ_g, of the phase delay in equation 19.1 is the geometric path-length difference of the radio signal traveling to each of the two elements. It is a function of the direction of the incoming radio waves and changes rapidly with the diurnal motion of the earth.

$$\phi_g = 2\pi\nu \left(\frac{\mathbf{b}\cdot\mathbf{s}}{\mathrm{c}}\right) \equiv 2\pi\nu\tau_g\,, \tag{19.3}$$

where $\mathbf{b}$ is the separation of the phase centers of the two antennas, $\mathbf{s}$ is a unit vector towards the centroid of the source, c is the speed of light and ν is the average frequency of the signal. The geometric time delay τ_g is also commonly used.

The total phase delay can be written as

$$\phi_T = 2\pi\nu(\tau_g + \tau_n) + \phi_d + \phi_V\,, \tag{19.4}$$

where it is convenient to define three other parts of the phase delay. First, any additional part of the phase delay which is also a linear function of frequency (like the geometric delay) is associated with τ_n, the non-dispersive time delay. It is so named because a pulse which travels through a non-dispersive delay medium will not increase in width. Two examples are tropospheric refraction and clock (timing) errors. That part of the phase delay which is dispersive is given by the pure phase term ϕ_d and two notable examples are the phase delay produced by the ionosphere and any frequency-independent phase changes in the antenna system. Finally, the visibility phase of the source, ϕ_V, depends on the source shape.

In summary, the total phase delay consists of: the geometric phase delay which is associated with the location of the source; the visibility phase which is associated with the internal structure of the source; and two contaminants, a dispersive and non-dispersive delay.

19.2.2 The Model Phase Delay

In nearly all interferometric observations, the approximate total phase delay can be estimated from the observing parameters, the monitored state of the array and antenna system, the weather conditions and other information. This inferred phase delay is called the *model phase delay* and is denoted by ϕ_M. The difference between the total phase delay and the model phase delay is the *residual phase delay* and the relationship is

$$\phi_T = \phi_M + \phi_R\,. \tag{19.5}$$

With an accurate model, the residual phase delay is small and relatively constant with time and direction (different sources).

The largest and best-determined part of the model is the *a priori geometric phase delay*, ϕ_a.

$$\begin{aligned}\phi_a &= 2\pi\nu_0\left(\frac{\mathbf{b}_0\cdot\mathbf{s}_0}{c}\right)\\ &= 2\pi\nu_0(L_x\cos\delta\cos H - L_y\cos\delta\sin H + L_z\sin\delta)\,,\end{aligned} \tag{19.6}$$

where

$$\begin{aligned}\mathbf{b}_0 &= \hat{\imath}L_x + \hat{\jmath}L_y + \hat{k}L_z\,,\\ \mathbf{s}_0 &= \hat{\imath}\cos H\cos\delta - \hat{\jmath}\sin H\cos\delta + \hat{k}\cos\delta\end{aligned} \tag{19.7}$$

are the a priori antenna-pair separation and source direction, respectively.[1] ν_0 is the a priori average frequency. The baseline components in equation 19.6 are in units of light travel time, although they can be expressed in linear units or in wavelengths with a simple scaling term. The declination δ and hour angle H are the source coordinates.

Any additional terms in the model phase delay cover a variety of *known* causes of phase delay which affect the fringes. These include changes within the antenna system (fringe rotation, parallactic angle effects, phase quadrature circuits, cable length monitoring, phase calibration injection signals), and more external changes (tropospheric and ionospheric refraction, antenna structure distortions).

19.2.3 Phase (Lobe) Ambiguities

There is one major complication in the determination of the total phase delay from the known model delay and the measured residual delay. Since $r(t)$ is nearly a perfect sinusoid with time, the residual phase delay ϕ_R can only be measured between $-\pi$ and $+\pi$. This phase is denoted by the fringe phase, ϕ_f,

$$\phi_f = \phi_R \mod 2\pi\,. \tag{19.8}$$

[1] $\hat{\imath}$ points to $\delta = 0, H = 0$; $\hat{\jmath}$ points to $\delta = 0, H = -\pi/2$ (East), $\hat{k}$ points to $\delta = \pi/2$ (NCP).

Thus,

$$\phi_T = \phi_M + \phi_f + 2\pi n\,, \tag{19.9}$$

where n is any integer, positive, negative, or zero. The arbitrary addition of multiples of 2π in the phase is called *lobe ambiguities.* We generally call a set of interferometric observations *coherent* if $n = 0$, and *non-coherent* if $|n| \gg 0$. For low values of $|n|$, there is hope of reaching phase stability (coherence) and methods will be discussed below. For example, the determination of the location of a point source shown in figure 19.1*b* is still possible even with phase ambiguities because we have lots of observations of the source. Often, a quick inspection of the set of fringe phases from observations of many sources over time will indicate the lack of coherence. If the phase residual is randomly scattered over 2π radians, there is little coherence. If the phase residual is relatively constant over direction and time to within about 1.5 radians, there is coherence.

For radio arrays with baselines less than about 100 km (except at frequencies above 20 GHz), lobe ambiguities are uncommon because the model phase delay can be accurately determined to less than a radian of phase. For most VLBI arrays with baselines greater than 500 km, model errors in the a priori parameters and refraction in the atmosphere and ionosphere produce phase changes much larger than one revolution in a matter of minutes of time and a few degrees of sky. However, n is generally small and phase stability can sometimes be obtained. For most optical arrays n is large and true phase stability is difficult to obtain.

But, all is not lost with VLBI. First, if the source signal is sufficiently strong (see section 3.8 on page 53 for this criterion), closure phase and closure amplitude can be used to image the source. Most of the common, large phase errors are antenna-based and cancel when closure quantities are taken. Phase ambiguities are, then, usually not a problem. High-dynamic range images can be made, but with the loss of absolute positional information. Weak sources cannot be imaged, however, because the closure quantities are too noisy and the source closure visibility phase cannot be determined.

Secondly, accurate astrometric information can be obtained, not from the total phase delay with its ambiguities, but from the rate of change of phase with time and with frequency. These rates are also subject to lobe ambiguities only if the data are sampled so sparsely in frequency or in time that the phase changes by more than one lobe between samples. The use of these phase rates, rather than the phase directly, decreases the astrometric sensitivity of the observations, but avoids the problem of lobe ambiguities.

19.3 Absolute Astrometry

In this section we will start with the fundamental equation of astrometry which is a simple interpretation of the geometric phase delay and discuss its solutions. We will then generalize by considering multi-element arrays. The next three subsections discuss the two major types of instrumental delay components and the method of removing the ionospheric component of the delay using several widely-spaced frequencies. The concluding subsection describes the celestial and terrestrial reference frames and the complicated motion of the earth.

19.3.1 The Fundamental Astrometric Equation

A residual phase delay will be obtained if the antenna-pair separation and the coordinates of the source used for the model phase calculation are in error, i.e. $\phi_R = \phi_g - \phi_a$. Using Cartesian coordinates defined in equation 19.6, we obtain

$$\begin{aligned} \phi_R &= 2\pi\nu\ (A\ \cos H + B\ \sin H + C) + \phi(\nu, t, \mathbf{s}) \\ A &= \Delta L_x \cos\delta + \Delta\alpha L_y \cos\delta - \Delta\delta L_x \sin\delta \\ B &= -\Delta L_y \cos\delta + \Delta\alpha L_x \cos\delta + \Delta\delta L_y \sin\delta \\ C &= \Delta L_z \sin\delta + \Delta\delta L_z \cos\delta + \Delta T \\ &\quad + \text{second order terms}\,, \end{aligned} \tag{19.10}$$

where the Δ denotes the difference (true$-$model). We have added a catch-all phase term, $\phi(\nu, t, \mathbf{s})$, which incorporates all non-geometric residual phase terms, given in equation 19.4, as a function of frequency, time and source direction. The ΔT in the C term corresponds to a clock difference between the telescopes. This term is important for VLBI observations in which independent clocks are used; it is insignificant for connected arrays. We have used seconds as the unit for the clock and the baseline lengths.

Least-squares solutions of the above equation can be obtained if the non-geometric phase term $\phi(\nu, t, \mathbf{s})$ is nearly constant over the observation interval and there are no lobe ambiguities in the residual phase, i.e. $\phi_R = \phi_f$. These conditions are usually met in kilometer-sized arrays. The residual phase then varies sinusoidally with a period of 24 sidereal hours with an offset term. The amplitude of the sinusoidal part contains information about $\Delta L_x, \Delta L_y, \Delta\alpha, \Delta\delta$, while the phase offset, C, depends only on $\Delta L_z, \Delta\delta, \phi(\nu, t, \mathbf{s}), \Delta T$.

Two practical solutions of equation 19.10 are given by Elsmore and Mackay (1969) and Wade (1970). In the first experiment, 78 sources were observed for 12 hours, and A, B, C were determined for each. A typical plot of the residual phase is shown in figure 19.1*a* and *b* for a source with a small and a large positional error. Note that the phases in figure 19.1*b* do contain several lobe ambiguities; however, the continuous monitoring of the source makes the phase connection obvious. Had the source fringe phase been sampled only a few times during the 12 hours, a solution for A, B, and C may not have been possible.

In the Wade experiment, about 19 sources were observed three times in each 24-hour observation run. This experiment was more efficient (56 sources in 72 hours observing, versus 78 sources in 468 hours) because continuous observations of a source are not needed to determine A, B, C, if the position errors

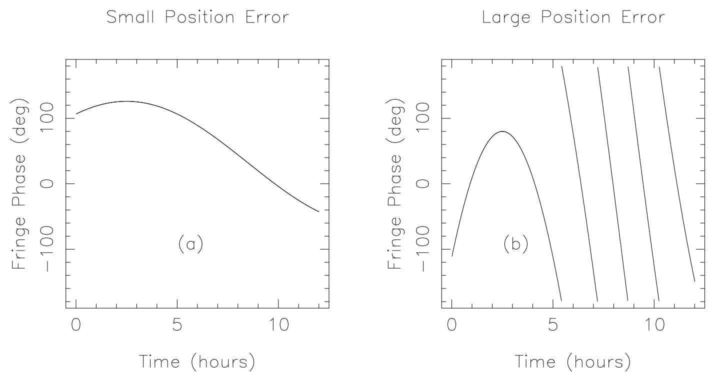

Figure 19.1: (*a*) A small position error. (*b*) A large position error which produces lobe ambiguities.

are small and lobe ambiguities do not occur. The technique of cycling amongst many sources, each observed several times, is still generally used as the basic astrometric observing schedule.

Both groups realized that *absolute* declinations could be determined from the observations; that is, the derived declinations did not depend on the assumed positions of the radio sources. Once $\Delta L_z, \phi, \Delta T$ (assumed constant) were determined, $\Delta\delta$ was fully defined. This is clearly related to the fact that the extreme range of the geometric phase delay depends only on $\cos\delta$.

On the other hand, absolute right ascensions cannot be determined. For example, manipulation of the A and B terms shows that a rotation of the antenna separation around $\hat{k}$ (the orientation of the equatorial component of the baseline) cannot be distinguished from a change of α in the opposite sense. Furthermore, in equation 19.10 we assumed that

$$\alpha = t_s - H\ , \tag{19.11}$$

and that $\Delta\alpha = -\Delta H$; where the local sidereal time is t_s. Clearly a time offset cannot be distinguished from a right ascension offset in all of the sources; hence $\Delta\alpha$ could be replaced by $\Delta t_s - \Delta\alpha$ in equation 19.10. Thus, the right ascensions measured by arrays are relative to a definition of the orientation of the equatorial component of the baseline and the origin of the system of time. The realization of what quantities are absolutely determined and what quantities are determined with respect to other parameters becomes more subtle and important in the complicated situations which will be discussed below.

As with all measurement sets in which many parameters are to be determined, the number of independent measurements must be no less than the number of unknown parameters. The further increase in the number of independent measurements makes the solution more robust and less influenced by outliers. For a two-element array observing s radio sources, the number of unknowns is

$(3+2s)$; 3 baseline coordinates, $2s-1$ source coordinates, and one phase offset. The number of independent measurements can not be larger than $3s$ regardless of the number of observations, since only three parameters are associated with a source, A, B, C. It takes three observations of the source to determine them, but more observations only increase their accuracy. Therefore, at least three sources are necessary to determine a solution. With a large number of sources, less than three observations of each are needed to obtain a well-defined solution.

19.3.2 Multi-element Arrays

The number of independent measurements for an $n-$element array is $(3s(n-1)n/2)$, and increases more quickly than the number of unknowns, which is $(2s-1+4(n-1))$. However, when observing strong point sources with a well-behaved instrument, the fringe phases are *antenna-based,* meaning that

$$\phi_{i,j} + \phi_{j,k} + \phi_{k,i} = 0\,, \tag{19.12}$$

where $\phi_{i,j}$ is the residual phase between any two elements (i,j). In this case the value of the third phase in antenna triangle is known from the other two phases and we often say that the phases *close* over the triangle. So, for a multi-element array with closure in the fringe phases, [2] there are really just $3s(n-1)$ independent measurements. All the information, then, can be obtained by using only antenna-pairs with respect to one antenna—usually called the *reference* antenna. The fringe phase for any other baseline is a simple sum of the relevant baselines with respect to the reference antenna and contain no additional information.

In practice all baselines are used in the solution anyway. First, if the signal-to-random-noise ratio over an observation is not large (typically < 5 which is a relatively weak source), some accuracy is gained by including all baselines. More important, a solution with all baselines is a good check on the quality of the data and the solution. The most important and common non-closure phase is ϕ_V, the visibility phase. If the source structure can be determined from the observations (with self-calibration techniques) or externally, the ϕ_V term can then be moved from the residual phase to the model phase. If it remains unknown, then the use of all $n(n-1)/2$ baselines in an array of n elements is superior to just using the $(n-1)$ baselines to the reference antenna. The variations in the ϕ_V will cancel a bit better with more baselines.

19.3.3 Common Non-dispersive Phase-Delay Terms

All of the astrometry is contained in the solution to equation 19.10. However, the simple illustration in the previous section assumed that $\Delta T + \phi(\nu, t, \mathbf{s})$ was a constant over the observation period. That is, when the best model phase delay

[2] Only the fringe phase must close. The model may contain non-closing phases which remove their effects from the total phase.

was removed from the total observed phase delay, the resultant fringe phase contained only the residual geometric phase delay plus a constant phase for each antenna. This is not the case for VLBI experiments where large unknown variations of phase wreak havoc in trying to obtain good astrometric solutions. We can only give a superficial treatment of some of the more important terms.

Any non-dispersive phase delay terms, like the geometric delay, scale with frequency. Thus, there is no simple multi-frequency experiment which can separate the two components by their frequency dependence. Of course many non-dispersive components have a different temporal and directional dependence than the geometric delay and can be "filtered" by proper averaging. Nevertheless, there are phase delay terms which tend to have a 24-hour period and can mimic the time dependence of the geometric delay. The discussion of a software system, SOLVE, to analyze such data is given in section 19.4.7.

19.3.3.1 Tropospheric Refraction

Tropospheric refraction generally produces the largest uncertainty in the measurement of the total phase delay. It has two major components: the dry and wet terms. Because of day-to-night differences this phase delay can display a 24-hour period. The parameters which measure the total delay are the zenith path delays, τ_{dry} and τ_{wet}. The dry air component is measured by the total pressure (i.e. the amount of atmosphere above the telescope) and is approximately; see Thompson, Moran and Swenson (1986); (hereafter TMS), page 411).

$$\tau_{dry}\,[\mathrm{cm}] \approx 0.228 P_{tot}\,[\mathrm{mbar}]\,, \tag{19.13}$$

while the wet delay τ_{wet} is

$$\tau_{wet}\,[\mathrm{cm}] \approx P_{wv}\,[\mathrm{mbar}] \approx 6.3w\,, \tag{19.14}$$

where P_{wv} is the partial pressure of the water vapor and w is the effective column size of water above the telescope. The dry delay at normal pressure is about 230 cm (7 ns); the wet delay varies from 1 to 30 cm (0.03 to 1 ns). The actual phase delay depends on the zenith angle, z, and various zenith angle dependencies, called *mapping functions,* $m(z)$, are used. For a plane-parallel atmosphere approximation, $m(z) = \sec z$. More complicated models are now in use (see section 9.6.1 on page 178 for further discussion).

Although the above formulation models the gross path-delay difference, the troposphere is characterized by micro-turbulence at size scales from meters to kilometers, with little effect on the ground-based measurements. It is believed that the irregularities in the distribution of water vapor dominate the phase irregularities, which occur with characteristic time scales between seconds and hours. The range of delay fluctuations are approximately 0.1 to 1.0 cm and the correlation scale ranges between meters to many kilometers (see sections 9.6 and 9.6.1 for further discussion of atmospheric phase errors).

Water vapor radiometers (WVR) have been used to measure the delay of the wet component by measuring the water vapor emission at several frequencies in the direction of the radio source. The results are unreliable at accuracies of better than about 1 cm, and the effects of liquid water are deleterious to this measurement (TMS, page 436).

19.3.3.2 Clock Errors

The time delay of the signal from the radio source at any antenna is measured with respect to a telescope (station) clock. If one centralized clock is used and its pulses are distributed among the telescopes, then telescope clock errors will depend on properties of the propagation of the clock data. These errors are typically about < 0.001 ns.

If independent clocks (for most VLBI observations) are used, then their relative offset, drift, and acceleration will produce concomitant time delay differences between the antennas. A typical offset in the independent setting of two clocks is about 1000 ns and a reasonable rate difference is 1 part in 10^{14}. At a frequency of 10^{10} Hz, the difference in phase can accumulate to a radian in about 300 seconds (TMS, page 282).

19.3.3.3 Instrumental Path Lengths

The radio signals are transported from the feed to the correlator along many types of signal paths. Most of the path-length changes are non-dispersive and are produced by temperature induced changes in the antenna system or physical motion of the telescope signal path when tracking the diurnal motion of a source. These path-length differences can be measured electronically by sending signals along the various paths and measuring changes with time (usually by reflecting the signal back to the original location). Thus, most of the local path-length changes can be monitored and included in the model phase delay (Thompson and Bagri 1991).

19.3.4 Common Dispersive Phase Delay Terms

19.3.4.1 Ionospheric Phase Delay

The phase delay through the ionized plasma of the earth's ionosphere and the solar corona varies as ν^{-1} at frequencies above the plasma resonance frequency, generally a few megahertz. The major ionospheric phase delay occurs at an altitude of 200 to 500 km and contains a general day/night component plus ionospheric events which are probably associated with the wave propagation in the ionic medium caused by solar storms and by the day/night transition. The amount of phase delay depends on the electron density and the magnetic field strength along the line of sight. Radiosonde measurements of the electron content along the line of sight can be useful in determining a reasonable model of the ionospheric refraction; however, the shorter term events are unpredictable.

A typical value of the time delay of the quiescent component of the ionosphere, τ_{ion}, during the day is

$$\begin{aligned} \tau_{ion}\,[\mathrm{cm}] &\approx 2400\,\nu^{-2}\,[\mathrm{GHz}] \\ \phi_{ion}\,[\mathrm{rad/m}] &\approx 1700\,, \end{aligned} \tag{19.15}$$

or about 80 ns at 1 GHz (TMS, page 445). ϕ_{ion}[rad/m] is the phase shift in radians per meter of wavelength. The delay is about 200 cm (6 ns) during the night, about that from the entire tropospheric component. Because of the large scale height of the ionosphere during quiescence, the difference in time delay between telescopes is

$$\delta\tau_{ion} \approx \tau_{ion}\left[1-\exp\left(-\frac{c\tau_g}{500\,[\mathrm{km}]}\right)\right]\,, \tag{19.16}$$

with no strong dependence on zenith angle until $z > 70°$.

19.3.4.2 Visibility Phase

The visibility phase associated with the internal structure of the source is dispersive. It, with the visibility amplitude, are used to construct an image of the source. For astrometric observations of small-diameter sources where it is assumed that the source is point-like, the visibility phase is a noise-like term. Fortunately, it is never larger in magnitude than π radians and, thus, never larger than a delay in ns of $\frac{\pm 0.5}{\nu[\mathrm{GHz}]}$.

19.3.4.3 Phase Changes

Dispersive phase changes with time in array systems are detrimental to accurate astrometric results. As we shall see in dealing with the group delays, such phase changes must be removed before comparison of the phase between frequency channels or IF's can be useful.

19.3.5 Multi-Frequency Observations

Since various phase delays have different frequency behavior, they can be separated by observing a radio source at several frequencies. Consider the residual phase delay measured for simultaneous observations of a source at two frequencies ν_1 and ν_2. After removing the best phase model at each frequency we are left with

$$\phi_1 \equiv 2\pi\nu_1\tau_1 = 2\pi\nu_1\tau_0 + I/\nu_1 \tag{19.17}$$

$$\phi_2 \equiv 2\pi\nu_2\tau_2 = 2\pi\nu_2\tau_0 + I/\nu_2\,, \tag{19.18}$$

where $\phi_{1,2}$ are the measured fringe phase delay at frequencies $\nu_{1,2}$ and $\tau_{1,2}$ are the equivalent measured residual time delays. The sum of the geometric and tropospheric time delay is τ_0 and is frequency independent, by definition; I is the ionospheric phase delay parameter. Other additional residual phase delay terms have not been included in the above terms. By taking the appropriate sums of equations 19.17 and 19.18 we can obtain solutions for τ_0 and I

$$\tau_0 = \frac{\nu_1^2\tau_1 - \nu_2^2\tau_2}{\nu_1^2 - \nu_2^2} \tag{19.19}$$

$$I = \frac{2\pi(\tau_1 - \tau_2)}{1/\nu_1^2 - 1/\nu_2^2}\,. \tag{19.20}$$

19.3.6 Reference Frames and the Motion of the Earth

The fundamental astrometric equation 19.10 is somewhat vague about reference frames and coordinate systems. Because radio arrays are located on the earth which has a variety of motions and surface deformations, the geometric delay of the signal from a source in the sky between two telescopes is a complex function of time. Some of the major effects are discussed below. The geometry of the terms have been worked out and are incorporated in the SOLVE and CALC algorithms which are discussed in the next section.

19.3.6.1 Celestial Coordinate Frame

A satisfactory coordinate frame has its origin at the solar system barycenter because it is inertial, i.e., non-rotating (except for galactic rotation). The z-axis is defined in the direction of spin-axis of the earth at a specific epoch and the (x, y) axis are defined by the direction of the vernal equinox at the same epoch. Since the extragalactic radio sources form a convenient grid of stationary objects in the sky, their direction with respect to this coordinate frame must remain fixed. At the present time the coordinates of 23 primary radio sources are used (Arias, Feissel and Lestrade 1988; Ma et al. 1990).

There are two systems in general use; one called B1950.0 and the other J2000.0. The B1950.0 system is tied to star catalogs on the FK4 system, those on J2000.0 are associated with the FK5 star catalogs. Besides the equinox difference in the definition of these two coordinate systems, there is a slight change of the spin-axis with time between them. Hence, the conversion of coordinates from B1950.0 to J2000.0 requires an *effective* date (epoch) of the observation. For example, all further B1950.0 positions used for the VLBA and VLA have an effective epoch of 1979.5, the date of the first astrometric observations of primary calibrators at the VLA. All further positions derived at the VLA, regardless of the time of the observations, are tied to the positions of these primary calibrators, hence this same epoch (Donat and Boksenberg 1994, section L).

19.3.6.2 Terrestrial Coordinate Frame

The origin of the terrestrial coordinate frame (ITRF) is the center of mass of the earth. The coordinate axes, however, are difficult to define. Because the earth is not a rigid body, any axes definition tied to surface objects would distort with time. For this reason, the origin and axes directions are implicitly defined by the coordinates adopted for about 50 terrestrial sites. These coordinates and their rate of change are adjusted so that there is no net rotation of the crust of the earth. Many of the fiducial points are well-known radio telescopes. This grid does define the location of the center of mass to an accuracy 10 cm, and there is continuity at the $< 0\farcs03$ of the terrestrial pole and reference meridian (Greenwich) previously defined (Boucher and Altamimi 1989).

19.3.6.3 Earth Orientation

The orientation in space of the spin axis of the earth precesses because of the gravitational pull of the sun, moon and planets on the equatorial bulge of the earth. The major term, the annual precession caused by the sun and the moon, is a precession of the pole with respect to the plane of the earth orbit (ecliptic). The angle of inclination is about $23.5°$ with a precession period of 26000 yr. Planetary gravitational force causes additional precession, called nutation, which decreases the inclination of the ecliptic by about $50''$ per century. Since the z-axis of the fundamental celestial coordinate system is parallel to the orientation of the spin-axis at the equinox date, the motion of the pole must be carefully tracked to keep the extragalactic source coordinates from apparently rotating with time.

In 1980 the IAU adopted precession/nutation models which are to be applied to keep the extragalactic source positions constant (Kinoshita 1977). However, it was soon found with VLBI observations, that the model was systematically in error and that additional components of periods of 18.6 yr, 1.0 yr, 0.5 yr and 14 days are present at the level of $0\farcs01$ (Herring, Gwinn and Shapiro 1986). The two relevant nutation terms are $d\psi$ (longitude), the change of the right ascension and $d\epsilon$ (latitude), the change of declination. These additional terms are predicted by the IERS standard and are generally included in the initial model for the phase delay for consistency among different experiments.

19.3.6.4 Polar Motion

The intercept of the spin axis of the earth is not rigidly tied to the earth; it moves up to 15 m on the earth surface, about $0\farcs5$ in direction. The main component, called *Chandler wobble,* has a period of about 1.2 yr and is caused by the season change in the mass of the oceans and the atmosphere, mixed with lower frequency terms. The zero point of the pole is taken, arbitrarily, as the mean pole location between 1900–1905.

19.3.6.5 Time and Earth Rotation

The rotation of the earth with respect to the extragalactic sources defines sidereal time. However, time based on the length of a solar day, *universal time (UT),* is a more useful timekeeping quantity. The relationship between UT and *GMST (Greenwich mean sidereal time)* is defined by a mathematical formula (Donat and Boksenberg 1994, L2). With no correction for the polar motion, the time is called UT0 and it varies slightly with location on the earth. UT1 is the time system with polar motion removed and it is the same for all locations on the earth. But, UT1 and GMST include variations in the rotation of the earth, so they do not run uniformly. However, some of the known changes in the rotation of the earth (annual and semi-annual terms) are included in the time system UT2, which is a more constant clock. Anomalous variations of the length of the day by a few milliseconds are not uncommon and are unpredictable. They can be measured with sensitive observations. They are probably caused by storms or global wind patterns which cause an exchange of angular momentum between the earth and the atmosphere.

An external uniform time system is called *TAI (international atomic time).* It is defined from standard oscillators located at several institutions around the world. This time is based on a *standard second* which is defined as a large number of periods of the radiation between two hyper-fine ground states of the Ce^{133} atom. The time UTC (Coordinated) is tied to TAI but differs by an integral number of seconds; hence it is a uniform time system. However, leap seconds are added to UTC once or twice a year, as needed, in order to keep the difference between UTC and UT1 less than 0.9 sec. The most common time systems used in VLBI observations are UT1 (closely tied to sidereal time) and TAI or UTC, which are nearly uniform. There is also TDB, barycentric dynamical time, which differs from TAI by variations in the gravitation potential around the earth's orbit (Seidelmann 1992).

19.3.6.6 Plate Tectonics

It is now known that the earth's crust consists of large plates, floating on the mantel, that move apart from each other by as much as 10 cm yr^{-1}. Thus, station locations will also change over time. All of the previous effects change only the apparent orientation between two antennas, but plate tectonics will also change their separation. Diurnal and semi-diurnal earth tides and ocean loading (the depression near the coast line at high tide because of the extra weight of the water) are important at the mm level. These geodetic changes are discussed in chapter 18 in more detail.

19.3.6.7 Earth Orbital Parameters

The motion of the earth around the sun produces aberration in the position of a star of up to $20.''5$ when the velocity of the earth is perpendicular to the direction of the star. Thus, the orbital motion of the earth must be known to correct the observations to the barycenter origin of the celestial coordinate system. The orbit of the earth depends on the location and mass of the solar system objects, and these have been determined over many years from optical transits, radar ranging, spacecraft trajectories and lunar laser ranging. The results are contained in solar system ephemerides, and the most accurate is called the DE202 ephemeris, compiled by JPL. An independent check on the accuracy of the ephemeris would be a comparison of its coordinates of a solar system object with that derived from radio interferometry. However, the position of solar system radio objects are difficult to determine because of their large size (Cole 1976).

Pulsar timing positions are also tied to the orbit of the earth and thus rely on the solar system ephemeris. A detailed comparison of pulsar timing positions and interferometric positions would also (indirectly) be a comparison of the solar-system dynamical reference frame to that of quasars (Bartel et al. 1985; Gwinn et al. 1986; Fomalont et al. 1992). Another interesting comparison between pulsar timing and interferometric positions is related to the origin of right ascension, the vernal equinox. A radio array determines absolute declination and relative right ascensions; pulsar timing determines absolute ecliptic declination and relative ecliptic right ascensions. By comparing the positions of at least two pulsars using the two techniques, it is possible to compare the origin of right ascension of the two systems.

19.4 VLBI Observations

The major problem with astronomical observations using the VLBI technique was briefly mentioned in section 19.2.3: the observations are not strictly coherent because of lobe ambiguities between the measured fringe delay, ϕ_f, and the more fundamental (residual) phase delay, ϕ_R, which is used in the fundamental astrometric equation 19.10. The severity of the lobe ambiguity problem varies enormously with observing frequency, weather conditions, baseline length, clock stability and other instrumental characteristics of the array (cf. Rönnäng 1989).

19.4.1 Phase Derivatives with Time and Frequency

The observable derivatives of the fringe phase delay with time and frequency are much less subject to the lobe ambiguities between the residual phase and the fringe phase. These rates also contain useful astrometric information about the radio source and, although the results are not as accurate as that using the phase delay directly, the precision obtained with VLBI data over the last decade is impressive. The methods of analysis are given in this section.

The rate of change of phase with time is a natural observable from the data and it is the large size of these rates which cause the bulk of the lobe ambiguity problems. Figure 19.1*b* showed the fringe phase of a source which has a large position error and obvious lobe ambiguities. If a few short observations are taken of the source over the 12-hour period, there is little chance that the correct values A, B and C can be ascertained and useful astrometric information obtained. However, if each observation is sufficiently long, then the phase change with time can also be obtained. This extra information can be used to solve equation 19.10 to connect the phases unambiguously, even over several lobe rotations during the observational period.

Even if only the phase rate, $d\phi_f/dt$, is determined for an astrometric run of many sources, astrometric information can be derived. It is generally more useful to consider the residual delay rate,

$$d\tau_f/dt = \frac{1}{2\pi\nu}\frac{d\phi_f}{dt}\,. \tag{19.21}$$

Unless the phase is sampled so sparsely in time that lobe ambiguities occur between samples, the residual delay rate should have no ambiguities.

In a similar manner, if we observe a source simultaneously at several closely-spaced frequencies, the change of phase with frequency is also a measure of the phase delay without the normal ambiguity problems. The derivative of phase with frequency is called the *group delay* τ_{gr} and it is given by

$$\tau_{gr} = \frac{1}{2\pi\nu}\frac{d\phi_f}{d\nu}\,. \tag{19.22}$$

It turns out that with the use of the group delay, the solution to the astrometric equation 19.10 is very similar to that used for the phase directly.

19.4.2 Solution Using the Phase Rate

The measurement of the phase rate is straight-forward and does not require any more complication that a simple, one frequency array. Generally, the fringe phase is sampled at an interval of s sec and from even a short scan of no more than a few samples, a reasonable phase rate can be obtained (assuming sufficient signal-to-noise in the observation). However, there is a phase rate ambiguity of $1/s$ Hz where s is the sampling time in seconds. The fundamental astrometric equation 19.10, after taking the derivative with time, becomes

$$\begin{aligned} d\phi_R/dt &= dH/dt\ 2\pi\nu(-A\sin H + B\cos H + C) \\ A &= \Delta L_x\cos\delta + \Delta\alpha L_y\cos\delta - \Delta\delta L_x\sin\delta \\ B &= -\Delta L_y\cos\delta + \Delta\alpha L_x\cos\delta + \Delta\delta L_y\sin\delta \\ C &= d\Delta T/dt + d\phi(\nu, t, \mathbf{s})/dt \\ &\quad + \text{second order terms}\,. \end{aligned} \tag{19.23}$$

The term dH/dt is the sidereal rate of 7.27×10^{-5} rad s^{-1}. We have replaced ϕ_R with ϕ_f since lobe ambiguities should not be a problem with reasonable sampling of the phase in time.

Detailed discussion of the solution based on fringe rate are given elsewhere (e.g., by Wohlleben, Mattes and Krichbaum 1991, page 105). The offset term in the fringe rate equation (C parameter) only contains instrumental error terms; i.e., no baseline or source parameters. Thus, the number of independent measurements is only $2s$ whereas the number of unknowns are $2s+3$ (section 19.3.1). Either the position of one source must be assumed (therefore both α and δ are relative), or the baselines must be known. Also, the accuracy of the source declinations are reduced near the equator.

19.4.3 Solution Using the Group Delay

Using the definition of the group delay in the above section, we can derive from equation 19.4 the effect of the main components of the phase terms on the group delay.

$$\frac{1}{2\pi}\frac{d\phi_T}{d\nu} \equiv \tau_{gr} = \tau_g + \tau_n + \frac{1}{2\pi}\left(\frac{d\phi_d}{\nu} + \frac{d\phi_V}{2\pi}\right) . \tag{19.24}$$

Two major components of the group delay are the geometric delay and the non-dispersive delay—which include the clock offset and the tropospheric delay. The non-dispersive phase effects produce identical phase delay and group delay measurements. Dispersive delay effects are caused by the ionosphere, the source structure and peculiar phase differences between different frequency channels in the antenna receiver system. Of course the total group delay is not measured, but a residual group delay after removal of an accurate a priori model can be measured. In order to determine the group delay with no lobe ambiguities, the frequency-sampling of the experiment must be carefully chosen, and an example is given in the next section.

19.4.4 Measuring the Group Delay

Since the group delay is determined by taking the difference in the observed phase amongst several frequency channels, any dispersive residual instrumental phase difference amongst these channels must be small. Many ingenious phase calibration methods are used in VLBI arrays in order to measure such phase differences as accurately as possible, and remove them in the phase model (see chapter 10). Often, the accuracy in the measurement of the group delay is limited by phase offsets between the different frequency channels and IF's, and a significant part of the instrumental calibration attempts to determine these phase offsets. Otherwise, the comparison of the phase observed at different frequencies in order to obtain the group delay is useless.

Clearly, the accuracy of the group delay will depend on the amount of spanned bandwidth in the observations. The frequency coverage does not have to be continuous, but the gaps between frequencies must be arranged so that lobe ambiguities between frequencies do not occur. The technique of using much larger spanned bandwidth than instantaneous bandwidth is called *bandwidth synthesis*. With good planning, the instantaneous bandwidth can be less than 5% of the spanned bandwidth (e.g., Ma et al. 1986).

Most VLBI arrays produce data simultaneously at many frequencies, within several hierarchies, and an example below will illustrate the determination of the group delay in a complicated condition. First, the experiment may contain two separate simultaneous observing bands, each band passing through independent receiver systems. Second, there are independent, widely-spaced frequencies, which are called IF's (intermediate frequencies) within each observing band. Each of these come from the same data stream through the front-end amplifiers, but are split apart in base-band converters. Any dispersive phase changes between these IF's, particularly phase differences which vary with time, must be accurately measured. Finally, each IF is usually decomposed into a number of equally-spaced frequencies, called channels. These channels are a consequence of the correlator processing over a wide range of delay. Over the years astrometrists and geodecists have devised many band, IF and channel frequency allocation schemes in order to maximize the spanned bandwidth and minimize the instantaneous bandwidth, while trying to obtain unambiguous group delays.

An example of the step-by-step processing of data to determine the group delay from one scan (typically a few minutes in length) and one baseline for a VLB experiment is given in figure 19.2. The frequency coverage is for a VLBA observation taken on June 28, 1993. The observation frequencies are 1410, 1418, 1442, 1482, 1586, 1642, 1674, 1698 MHz and each IF has five channels at a separation of -2, -1, 0, 1, 2 MHz from the center frequencies (this is not exactly the VLBA observing configuration). We shall assume that any delay rate with time has been removed over the scan and that the plotted phases are the average for each frequency channel at some fiducial time near the scan middle.

A reasonable set of the observed fringe phases for this experiment is given in figure 19.2*a*. This is the phase after removal of the best a priori model phase (in the correlator and elsewhere). The 40 frequencies are grouped into the 8 IF's, each with five channels. There is a strong phase gradient in the five channels of each of the eight IF's, which indicates that there is a large residual value of group delay. The phase slope across each IF is called the *single-band* or *narrow-band* group delay and there can be a slight scatter amongst the narrow band delays. If the phase difference between each channel in the narrow band is more than one lobe, there is an ambiguity of the slope, in this case 1000 ns, corresponding to the 1 MHz channel separation.

If we could, in fact, measure the residual phase (without phase ambiguities) for this experiment, it would be that shown in figure 19.2*b*. Notice that the slopes (group delays) in the individual IF's align well with the overall group delay given by the eight IF's. However, without the coarse slope given by the narrow band delay, the determination of the wide-band delay, the phase slope over all of the IF's, would be difficult to obtain.

Figure 19.2*c* shows the fringe phase after removing the average of the eight

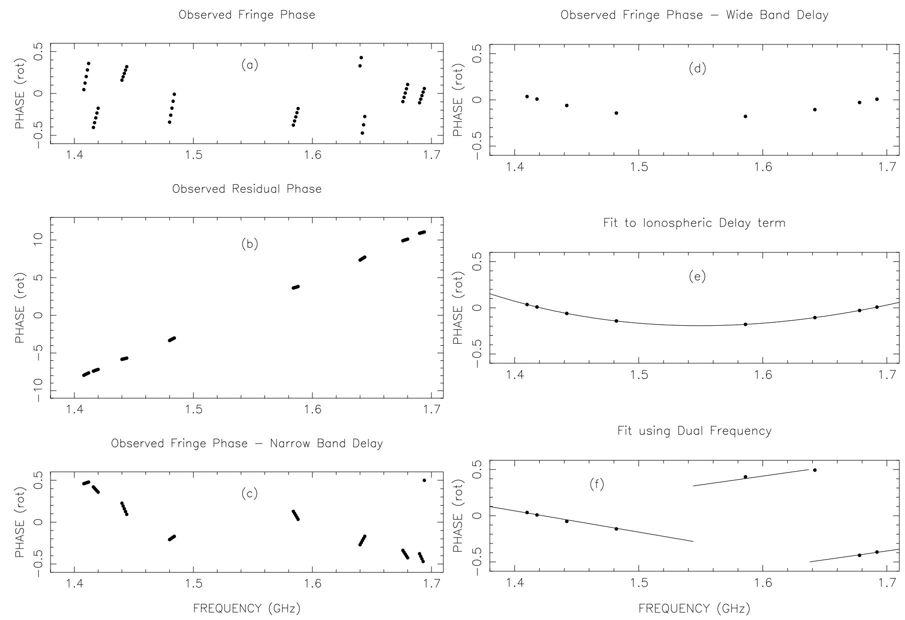

Figure 19.2: (*a*) The measured fringe phase for the 5 × 8 frequency samples. (*b*) The associated residual phase with no phase ambiguities for the same data. (*c*) The fringe phase after removal of the average narrow-band group delay. (*d*) The fringe phase after removal of the average wide-band group delay. The phase for the five channels in each IF have been averaged. (*e*) The fit to the fringe phase with an ionospheric term, as well as an offset and a delay slope. (*f*) The fit to the group delay in two sections (with an added phase offset) to determine the ionospheric content.

narrow-band group delays. Notice that even with the narrow-band correction, there are still some phase ambiguities across the frequency span. However, these can be sorted out by comparing the phases of the IF pair with the small frequency separation. In this particular case the 1410-1418 MHz pair defines the smallest separation with a resultant slope ambiguity of 125 ns.

The group delay which fits all of the IF phases is called the wide-band delay or the multi-band delay. After fits of the phases in figure 19.2*c* with the best group delay, the residual phases for each IF are shown in figure 19.2*d*. Only one phase is now plotted for each IF, since the channel phases average out to nearly zero near the center channel. Notice that there is a slight curvature remaining in the residual phases.

19.4.5 Measuring the Ionosphere Delay

The fringe phase curvature is indicative of a dispersive component of the phase. As seen in equation 19.15, the shape is indicative of an ionospheric component where the phase varies inversely with frequency. The most straight-forward solution is to fit the residual phases in figure 19.2*d* to a function of the form $\phi_0 + \tau(\nu - \nu_0) + I/(\nu - \nu_0)$ to determine the ionospheric component, the residual delay at some fiducial frequency ν_0, and the phase offset. Such a fit is shown in figure 19.2*e*.

In figure 19.2*f* we have illustrated the dual-frequency approach for the determination of the ionospheric delay component. To be more realistic, we have inserted a phase offset between the lower four frequencies and the higher four frequencies. Such a phase difference (presumably variable with time) could arise if the two groups of frequencies were associated with different receivers. There is no way that the phase can be connected between the two bands, and a continuous solution for all eight frequencies could not be obtained in this case. The difference in the group delays between the lower and higher frequencies is, of course, a measure of the curvature (ionosphere term) and the solution for the group delay and ionospheric delay are given in equations 19.19 and 19.20. This is the typical experimental technique for astrometric/geodetic observations with two groups of IF frequencies, one group at about 2.2 GHz and another at about 8.2 GHz.

19.4.6 Relative Accuracies of Phases and Delays

The relative errors in the various measurements of delay produced by random phase noise can be estimated from simple principles (see TMS, page 375 for a fuller treatment). Consider an observation of a point source for a short period of time, and measure the average phase, phase rate and phase derivative with frequency (group delay). Let $\Delta\phi$ be the error in the determination of the average phase, then the error in the determination of the time delay is

$$\begin{aligned} \text{using phase delay:}\quad & \Delta\tau \approx 0.2\Delta\phi/\nu \\ \text{using group delay:}\quad & \Delta\tau \approx 0.3\Delta\phi/\Delta\nu \\ \text{using fringe rate:}\quad & \Delta\tau \approx 50\Delta\phi/\left(\nu\Delta t\,[\mathrm{min}]\right)\ , \end{aligned} \tag{19.25}$$

where ν is the observing frequency, $\Delta\nu$ is the spanned bandwidth, and Δt is the observation length in minutes. Clearly, the error in the measured group delay is about $\Delta\nu/\nu$ of the error in the delay obtained directly from the fringe phase (if you can do it!). The length of an observation is generally limited to a few minutes because of curvature in the phase with time caused by second order clock errors between the two stations or by large atmospheric or ionospheric phase changes. With a typical scan length of five minutes, the error in the delay determined from the fringe rate is 20 times larger than the error in the measurement from the phase delay.

These relative sensitivities are misleading for normal VLBI observations. The *real* phase error in most cases is not strongly correlated with $\Delta\phi$, the phase error over the scan. There are other long term errors (atmosphere, ionosphere, baselines, clocks) which are generally non-dispersive and contribute $\Delta\tau$ to the errors in equation 19.25. For the phase delay and possibly the group delay, the more systematic errors are larger than the phase noise error.

19.4.7 Analyzing VLBI Data: SOLVE

The phase functions which are measured for each scan (a short observation) on each baseline are: the fringe phase, the residual group delay (both narrow band and wide band) and the residual delay rate; all at a time generally near the middle of the scan. These are associated with the model phase delay from which the total phase delay, group delay and delay rate can be found. Of course, as emphasized before, the total phase delay may be in error by many lobes, but the total group delay and total delay rate normally have no such ambiguities. While it is possible to use the total quantities in the subsequent analysis, it is more convenient to use the residual, smaller quantities. But the connection between the residual and the total values, through the model, must not be lost. Of course, many other parameters are needed to interpret the data: the time, the baseline, the weather conditions, the frequency, etc.

The most extensive set of software to edit and analyze VLBI data is called the Mark III Analysis Software. It has been developed over ten years at NASA Goddard, USNO, MIT and Harvard. It's goal is to derive better model parameters using the fundamental astrometric equation 19.10 with many more parameters inserted. The software description follows:

Data Handling: There is software for data input, storage, archiving, editing, plotting. Calibration from external observations can be added to the data base.

CALC: This part of the software calculates the model phase delay from a priori parameters decoupled from any calibrations which have been applied to the data (Robertson 1975). The input data are totals with no calibrations. This software also calculates the partial derivatives of the model with respect to any unknown parameter to be solved for such as the right ascension and declination of each source, the tropospheric zenith path delay and polar motions. The current version of CALC (version 8.1) includes many subtle contributions and the accuracy of the calculated delay is about 0.0001 ns.

SOLVE: SOLVE is the software package which uses a linearized least-squares method on the residual (phase, delay, delay rate) values in order to obtain improved parameter values. SOLVE uses the partial derivatives already calculated by CALC. For very large changes of some parameters, CALC should be rerun to recalculate the partials with a new model based on the improved value of these parameters in order to keep second order terms small. The atmospheric zenith path delay and the clock offsets at each antenna can be determined in a piece-wise continuous function of time, with some constraints in the size of the variation (Ma et al. 1986).

SOLVK: SOLVK is a variant of SOLVE developed by Herring at MIT using a Kalman filtering approach, instead of least-squares.

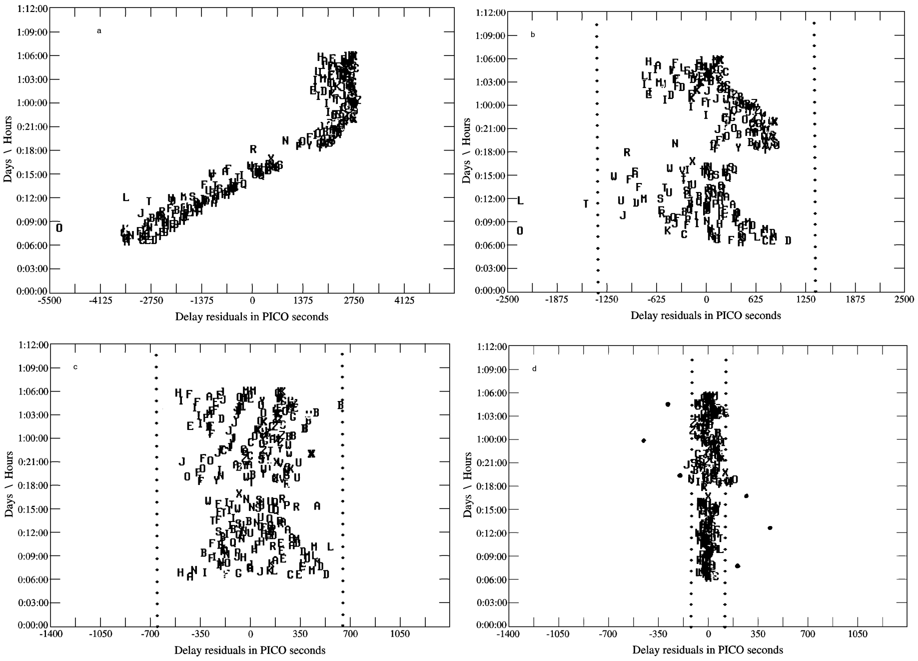

Figure 19.3: *(a, top left)* The residual group delay of the raw, edited data. The abscissa is the group delay with an offset of 954.8 ns removed. The ordinate is time, covering 24 hours of observations. *(b, top right)* The residual group delay after removal of the a clock offset, rate and acceleration *(see text.)* *(c, bottom left)* The residual group delay after removal of piecewise continuous clock and atmospheric contributions using three hour segments. See figure 19.4 for the piecewise continuous solutions. *(d, bottom right)* The residual group delay after correcting for nutation in longitude and latitude and the baseline separation of the two telescopes. The six additional points show the residual group delay for a source with a position error of $0.''03$.

A typical execution of the Mark III Analysis Software, SOLVE, will be described using figure 19.3. The figure shows the results from a VLBA run of five antennas (Kitt Peak, Fort Davis, Owens Valley, Fort Davis and Brewster) in August 1992. The data were observed at 8.2 GHz, with 400 MHz spanned bandwidth in four IF's, and simultaneously at 2.2 GHz, with 100 MHz spanned bandwidth in another four IF's. About thirty sources were observed over a 24 hours period, with each observation lasting about one minute of time.

The analysis will be shown for the observed group delay at 8.2 GHz. The ionospheric delay has already been removed using equation 19.19, in a manner similar to that shown in figure 19.2f. The residual group delays, at various stages of reduction, will be shown only for the Kitt Peak, Fort Davis (KP-FD) baseline.

The analysis usually proceeds in steps from the largest effects to the smallest effects, and a typical order is: (1) Removing outliers and bad observations;

(2) Using the narrow band group delay to sort out the correct ambiguity of the wide-band group delay (see fig. 19.2); (3) Determining the gross clock behavior; (4) Determining a piecewise continuous clock correction with time scale of a few hours; (5) Determining a piecewise continuous atmospheric correction with a time scale of a few hours; (6) Fitting for astrometric and geodetic parameters (nutation, source positions, UT1−UTC, polar motion, etc.).

An example of the reduction results follows:

Input: Using the SOLVE subroutines DBEDIT and DBCAL read in the data and any calibration values (weather parameters, phase calibrations etc.). Check the group delays and delay rates for outliers and edit appropriately. Phase cal calibrations should be applied to the residual data.

Clock Offset: Figure 19.3*a* shows a plot of the residual group delay for the KP-FD baseline (abscissa) as a function of time (ordinate) for each one-minute observation. The character type of the plotted point indicates the source. The time covers the entire 24 hours of the experiment. A known offset of 954.8 ns between the clocks was determined before correlating the data and has been removed before plotting the residuals. The clock offset between the two telescopes varies about 5 ns over the 24 hour period with an rms error around the mean of about 2250 ps.

The time, t, behavior is generally fit to the form $A + Bt + Ct^2$. For this data, SOLVE obtains $A = 3.4 \times 10^{-9}$sec$, B = -0.154 \times 10^{-12}, C = 0.079 \times 10^{-12}$ day^{-1}. The residual group delay, after removing this clock function, is shown in figure 19.3*b*, and the residual group delay rms for all sources is about 460 ps. Notice that this corresponds to about 3 lobes at 8.2 GHz; hence the use of fringe phase, rather than group delay, would have produced a nearly random scatter of fringe phase for this set of data.

Piecewise continuous clock and troposphere variations: The major parts of the residual in the group delay are usually hour to hour variations in the clocks at the two telescopes and in the atmospheres over the two sites. SOLVE permits a piece-wise continuous solution for both. A typical solution interval is about 3 hours. This parameterization clearly increases the number of unknown parameters, and can only be attempted with extensive observations, i.e. at least ten observations at a variety of elevations within 3 hours.

Figure 19.3*c* shows the residual group delay after parameterizing the clock and atmosphere over 3 hour intervals. The residual group delay is now only 218 ps. The clock and atmosphere delay solutions are shown in figure 19.4. The upper and lower envelope give the 1σ range for the solutions.

Astrometric Parameters: With reasonable clock and atmosphere parameters, turn on additional astrometric/geodetic parameters, in order of relative inaccuracy. For this particular experiment, the a priori positions of the radio sources were accurately known, but the location of the VLBA antennas were somewhat in error. In figure 19.3*d* the residual group delays, after solving for the location of the Fort Davis antenna (Kitt Peak was the reference) and nutation in longitude and latitude, are shown. The residual delay error has now been

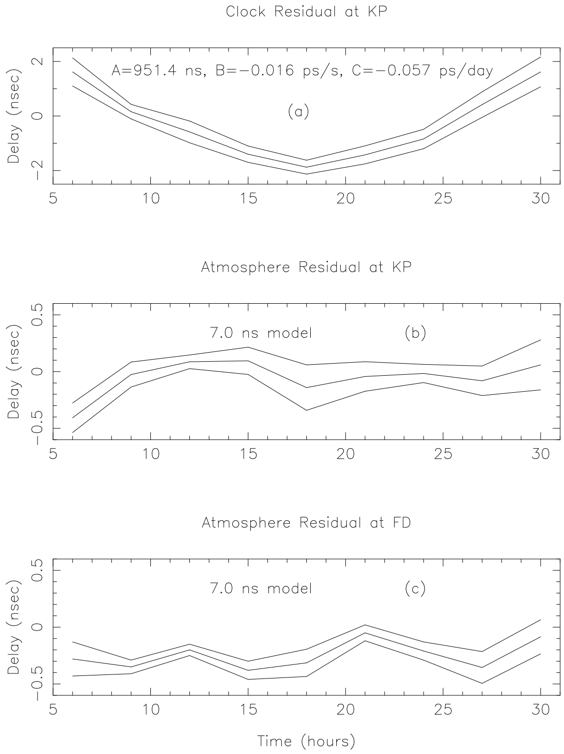

Figure 19.4: (*a*) The residual clock correction at Kitt Peak (Fort Davis is the reference station). This solution is in addition to the offset, rate and acceleration shown. (*b*) The residual atmospheric zenith delay derived for Kitt Peak at intervals of three hours. The 1-sigma error estimate is shown. The model used in CALC was a zenith delay of 7.0 ns. (*c*) The atmospheric solution derived for Fort Davis at intervals of three hours. The 1-sigma error estimate is shown. A zenith delay of 7.0 ns was used in CALC.

reduced to 39 ps, which corresponds to about 70° in phase at 8.2 GHz. In this plot we have added six fictitious observations of a source with a position error of $0\rlap{.}''03$. Over the KP-FD baseline of 30×10^6 wavelengths, such a position error produces a sinusoidal delay change of 500 ps amplitude over 24 hours.

The SOLVE software package contains a large number of parameters which describe earth orientation, plate tectonics, nutations of various periods, solar-system barycentric information, in addition to source positions and station locations. The robustness of solutions when attempting to determine a large number of parameters is difficult to ascertain because of the coupling of many of the parameters which have similar dependencies with time and angular dependence in the sky. However, this software is the best and most complete package to solve equation 19.10 available at the present time. Another powerful software package for the reduction of VLBI data has been developed by the JPL group. The original system, MASTERFIT (Fanselow 1983), has been improved and updated in the MODEST system (Sovers and Jacobs 1994).

19.5 Relative Astrometry

The previous sections dealt, mostly, with the spatial and temporal dependence of the phase and delay of sources observed over the sky. From these data, accurate source coordinates, station locations, and a host of other parameters could be derived. In order to separate out all of the parameters, complete sky coverage and rapid cycling between sources is important. However, if the main goal of the observations is to determine the coordinates or the motion of a radio source, then extensive observations over the entire sky are not necessary. Fortunately, the sky is filled with radio sources which are relatively small in diameter and stationary in the sky. Thus, the positions and motion of any radio source can be obtained by comparing its phase and delay with a nearby source, generally called a calibrator. Most observatories catalog a large list of calibrator sources which have been extensively observed, with good a priori coordinates and known small angular size.

There are two major types of comparative observations: *nodding* where observations are alternated in time between source and calibrator(s); and *simultaneous* where both source and calibrator(s) can be observed at the same time—generally because they are close enough in the sky to be in a single field of view of the instrument.

19.5.1 Nodding Observations

This type of calibration is by far the most commonly used at the VLA and will probably be nearly as common for the VLBA. This is the main technique that will be used for phase-connection of VLBA observations and is discussed in chapter 17. Here we will just briefly outline some of the astrometric problems which occur.

The observing technique simply alternates observations between a source and a calibrator (several calibrators will be discussed later) to measure the difference in their fringe phase delay as a function of time. If the cycling is fast and the sources are close in the sky, then all phase terms which are not associated with the residual geometric delay or structure phase of the sources should nearly cancel. However, tropospheric and/or ionospheric phase changes between the two sources in the sky, and between subsequent observations, may be sufficiently large so as to cause lobe ambiguities.

The plot in figure 19.5 illustrates the problems of VLBI nodding observations. A calibrator and source are alternately observed every 30 seconds, with the phase of each sampled every 6 second. If the phase of the calibrator observations can be connected, without lobe ambiguities, between scans, the observations are phase coherent and the nodding technique has been successful. The appropriate

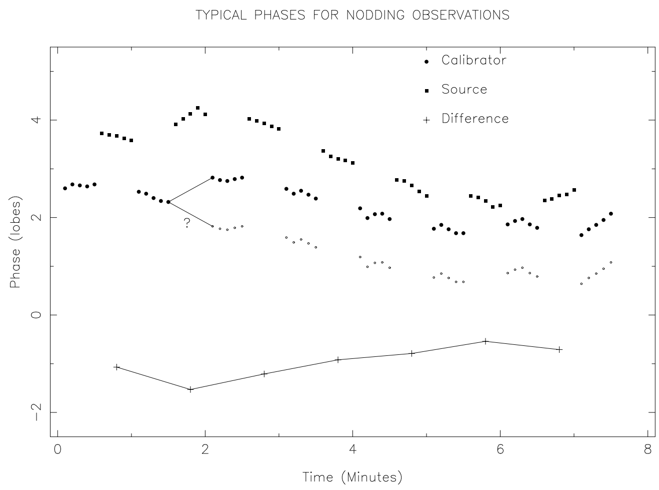

Figure 19.5: The calibrator phase observations are shown by the filled circles. A possible lobe ambiguity for the calibrator data is shown at $1.5 < t < 2.0$ sec. See text for its resolution. The source phase observations are shown by the filled squares. The phase difference is shown by the plus signs which are connected. This represents the calibrated phase of the source. An arbitrary constant has been added to the phase of the source and calibrator for clarity of presentation.

calibrated phase of the source is then the difference between the phase measured for the source minus the interpolated calibrator phase. This difference phase is shown by the lowest line in figure 19.5.

Notice that in the gap between 1.5 and 2.0 minutes, the interpolation of the calibrator phase is ambiguous with two possible paths. This, of course, will change the calibrated source phase. In this example, however, the source is strong and its phase can also be determined accurately every 6 seconds. The increase in the source phase during the calibrator gap between 1.5 and 2.0 minutes strongly suggests that the (unobserved) calibrator phase also increases during this period. The suggested interpolation would produce a smoother calibrated source phase. There are many indirect methods for resolving possible lobe ambiguities when observing with many frequencies and baselines, but sorting out lobe ambiguities is still more of an art than a science.

If fringe-phase connection cannot be obtained most of the time, then the group delays and delay rates can be used. However, the loss of accuracy is substantial.

The resultant fringe phase difference (or group delay) between the calibrator and the source should be dominated by a 24-hour sine wave, given by equation 19.10. To first order the derived position offset equals the residual positions of the source (true position minus assumed position) minus the residual position of the calibrator. For observations made at several epochs, the motion of the source can be determined with no assumption about the calibrator position, ex-

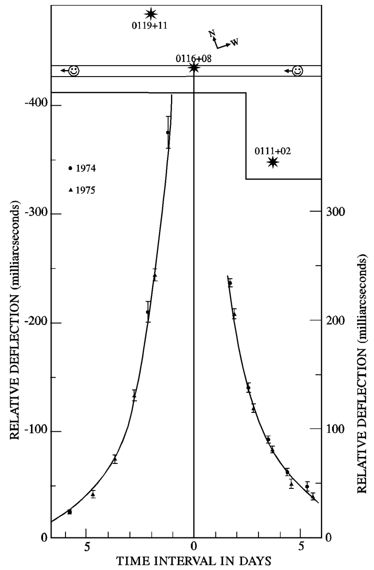

Figure 19.6: The measured change of position of 0116+08 as the sun passed in front of the radio source. This was derived from an astrometric analysis of the fringe phase over each day. The results are for 1974 and 1975 combined. The phase of 0116+08 was calibrated using the interpolated sum of the phases of 0119+08 and 0111+02 which lie on opposite sides of 0116+08.

cept that it is stationary in the sky. Several other effects can produce long term phase changes between the calibrator and source phase (e.g., station location errors, poor tropospheric models, and large clock errors), but these can be minimized with accurate astrometric observations in order to determine accurate a priori parameters.

19.5.2 Groups of Calibrators

The simple nodding experiment assumes that the temporal behavior of the fringe phase measured for the calibrator also applies to the source of interest, except for the fringe phase component caused by source position offsets and structure. However, there are many other phase errors which vary with both time and direction and produce phase differences between the source and calibrator with characteristic time scales of minutes to days. These phase errors generally vary directly with the source to calibrator separation in the sky and can be minimized using close pairs.

Observations of several calibrators surrounding a source of interest can determine both the temporal and angular dependence of the fringe phase. Two examples of the use of multiple calibrators are described by Fomalont and Sramek (1977) using the fringe phase and by Treuhaft and Lanyi (1987) using the group delay. Both experiments measured the gravitational bending of the radio waves past a massive body. The first experiment with a 35-km baseline observed a source, with two calibrators on either side over a period of two weeks, as it passed by the sun. Using a simple linear interpolation of the fringe phase observed for the two calibrators, the calibrated fringe phase at the source position could be determined (see fig. 19.6) using the interpolated fringe phase from 0119+11, 0111+02. Using an equation similar to equation 19.10, the relative position of 0116+08 was monitored as the sun passed near it, and a fit was made for the expected solar gravitational bending.

The second experiment used the NASA telescopes at Goldstone, California and in Australia, about 10000 km apart, five calibrators and the source were observed for six hours on two days separated by 12 days. On the first day, one source was near Jupiter. With the use of a large number of calibrators, the detailed dependence of the fringe phase with time and angle could be determined. Both experiments were dual frequency in order to remove the delay in the earth ionosphere and in the solar corona.

The use of several calibrators, instead of one calibrator, may not necessarily improve the accuracy of the experiment. First, the total integration time on the source will decrease and lessen the positional accuracy of the experiment. Second, the time gaps between observations of a calibrator will be significantly longer and lobe ambiguities will be more likely. Clearly, the proper design of a nodding experiment must take many factors into consideration.

19.6 Nodding Errors and Supportive Astrometric Observations

In order to lessen the phase errors in nodding experiments, periodic supporting astrometric observations may be useful. For example, in a period of thirty minutes, it is possible to observe about 20 strong sources located all over the sky in order to determine variable parameters associated with the array; such as, tropospheric zenith delay, clock offsets, phase calibration changes. With such observations every three hours, for example, the use of SOLVE can determine the variation of these important parameters which affect the phase. These updated parameter values for the model will decrease any systematic phase differences with the nodding observations.

19.7 Simultaneous Observations of Source and Calibrators

If two or more sources can be observed simultaneously, then phase connection between observation gaps is not a problem—there are no observation gaps. This desirable situation occurs if the source and calibrator(s) are closer than the size of the primary beam of the telescopes. The probability of this occurring depends on the density of sources in the sky, their angular size, the size of the telescopes

Table 19.1: Likelihood of background source in field

Freq. (MHz)	Resol. (mas)	Min. Flux (mJy)	Field Size (′)	Source> 7σ (%)
310	20	200	200	> 20?
610	10	200	100	10
1400	5	25	40	7
5000	2	25	13	2
8000	1.5	25	8	< 1

and the observing frequency. The likelihood that a sufficiently strong calibrator source will be in the primary beam of the VLBA antennas when pointing in a random direction is given in the table 19.1.

Of course, if you wish to determine the relative positions and motion amongst features of one extended radio source, the source angular size must be smaller than the primary beam of the telescopes.

Even when comparing the fringe phase of several sources within the primary beam of a VLBI observation, there may still be systematic phase errors between the sources, even with separations less than one degree. For example, the systematic elevation difference between objects in the primary beam can lead to systematic phase offsets if the tropospheric and ionospheric phase models are not very good. As with the case of nodding between several sources, if more than two sources are within the primary beam and are detectable, then angle dependent phase changes, as well as temporal changes, can be ascertained.

For further reading, these useful references are recommended: Babcock and Wilkins (1988b); Cornwell and Perley (1991); Davis and Booth (1993); Elsmore and Mackay (1969); Felli and Spencer (1989); Hirabayashi, Inoue and Kobayashi (1991); Mueller and Kolaczek (1993); Perley, Schwab and Bridle (1989); Reid and Moran (1988).

References

Arias, E. F., Feissel, M., & Lestrade, J.-F. 1988. "An extragalactic celestial reference frame consistent with the BIH terrestrial system (1987)". Annual report.

Babcock, A. K., & Wilkins, G. A 1988. *The Earth's Rotation and Reference Frames for Geodesy and Geodynamics*. Dordrecht: Kluwer.

Bartel, N., Ratner, M.I., Shapiro, I.I., Cappallo, R.J., Rogers, A.E.E., & Whitney, A.R. 1985. "Pulsar astrometry via VLBI". *Astron. J.* **90**, 319.

Biretta, J. A., & Owen, F. N. 1990. "Velocity structure of the M 87 jet: preliminary results". In J. A. Zensus and T. J. Pearson (Eds.), *Parsec-Scale Radio Jets*. Cambridge: Cambridge University Press, p. 125.

Boucher, C., & Altamimi, Z. 1989. "The initial IERS terrestrial reference frame coordinates". *Bull. Geodes.* **63**, 50.

Cole, T. W. 1976. "Absolute radio source right ascensions". *Observatory* **96**, 244.

Cornwell, T. J., & Perley, R. A. (Eds.) 1991. *Radio Interferometry: Theory, Techniques,* *and Applications*, Volume 19 of *ASP Conference Series*. San Francisco: Astron. Soc. of the Pacific.

Davis, R. J., & Booth, R. S. (Eds.) 1993. *Sub-Arcsecond Radio Astronomy*. Cambridge: Cambridge University Press.

Donat, III, W., & Boksenberg, A. 1994. *The Astronomical Almanac for 1994*. Washington: US Govt. Printing Office.

Elsmore, B., & Mackay, C. D. 1969. "The absolute determination of positions of 78 compact sources". *Monthly Notices Roy. Astron. Soc.* **146**, 361.

Fanselow, J. L. 1983. "Observation model and parameter partials for the JPL VLBI parameter estimation software MASTERFIT". Technical report.

Felli, M., & Spencer, R. (Eds.) 1989. *Very Long Baseline Interferometry*, Volume 283 of *NATO ASI Series C*. Dordrecht: Kluwer.

Fomalont, E. B., Goss, W. M., Lyne, A. G., Manchester, R. N., & Justtanont, K. 1992. "Positions and proper motions of pulsars". *Monthly Notices Roy. Astron. Soc.* **258**, 497.

Fomalont, E. B., & Sramek, R. A. 1977. "The deflection of radio waves by the Sun". *Comm. Astrophys.* **7**, 19.

Gwinn, L., Taylor, J. H., Weisberg, J. M., & Rawley, L. A. 1986. "Measurement of pulsar parallax by VLBI". *Astron. J.* **91**, 338.

Herring, T. A., Gwinn, C. R., & Shapiro, I. I. 1986. "Geodesy by radio interferometry: studies of the forced nutations of the Earth 1. Data analysis". *J. Geophys. Res.* **91**, 4745.

Hirabayashi, H., Inoue, M., & Kobayashi, H. (Eds.) 1991. *Frontiers of VLBI*. Tokyo: Universal Academy Press, Inc.

Kinoshita, H. 1977. "Theory of the rotation of the rigid earth". *Celes. Mech.* **15**, 277.

Ma, C., Clark, T. A., Ryan, J. W., Herring, T. A., Shapiro, I. I., Corey, B. E., Hinteregger, H. F., Rogers, A. E. E., Whitney, A. R., Knight, C. A., Lundquist, G. L., Shaffer, D. B., Vandenberg, N. R., Pigg, J. C., Schupler, B. R., & Rönnäng, B. O. 1986. "Radio source positions from VLBI". *Astron. J.* **92**, 1020.

Ma, C., Shaffer, D. B., De Vegt, C., Johnston, K. J., & Russel, J. L. 1990. "A radio optical reference frame. I. Precise radio source positions determined by Mark III VLBI: Observations from 1978 to 1988 and a tie to the FK5". *Astron. J.* **24**, 1284.

Mueller, I. I., & Kolaczek, B. (Eds.) 1993. *Developments in Astrometry and their Impact on Astrophysics and Geodynamics*. Dordrecht: Kluwer.

Perley, R. A., Schwab, F. R., & Bridle, A. H. (Eds.) 1989. *Synthesis Imaging in Radio Astronomy*. San Francisco: Astronomical Society of the Pacific.

Reid, M. J., & Moran, J. M. (Eds.) 1988. *IAU Symp. 129, The Impact of VLBI on Astrophysics and Geophysics*. Dordrecht: Kluwer.

Robertson, D. S. 1975. *Geodetic and Astrometric Measurements with Very Long Baseline Interferometry*. Ph. D. thesis, Massachussetts Inst. Techn. Reprint NASA TM X-922-77-228, Goddard Space Flight Center.

Rönnäng, B. O. 1989. "Geodesy-VLBI observables". See Felli and Spencer (1989), p. 298.

Seidelmann, P. K. 1992. *Explanatory Supplement to the Astronomical Almanac*. Mill Valley: University Science Books.

Sovers, O. J., & Jacobs, C. S. 1994. "Observation model and parameter partials for the JPL VLBI parameter estimation software "MODEST"—1994". Technical report.

Thompson, A. R., & Bagri, D. S. 1991. "A pulse calibration system for the VLBA". See Cornwell and Perley (1991).

Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 1986. *Interferometry and Synthesis in Radio Astronomy*. New York: Wiley-Interscience. First (1991) and second (1994) reprintings by Krieger Pub. Co., Malabar (Florida).

Treuhaft, R. N., & Lanyi, G. E. 1987. "The effect of the dynamic wet troposphere on radio interferometric measurements". *Radio Sci.* 22(2), 251.

Wade, C. M. 1970. "Precise positions of radio sources. I. Radio measurements". *Astrophys. J.* **162**, 381.

Wohlleben, R., Mattes, H., & Krichbaum, T. 1991. *Interferometry in Radioastronomy and Radar Techniques.* Dordrecht: Kluwer.