N79 21464

# Principles of Very-Long-Baseline Interferometry

Irwin I. Shapiro
Department of Earth and Planetary Sciences
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract. We present the basic principles of very-long-baseline interferometry as related to its use in the determination of vector baselines, polar motion, and earth rotation.

## Introduction

Ten years ago, almost to the day, the first successful bandwidth-synthesis VLBI measurements were made. It is thus appropriate to now review the principles underlying the technique. The review will be restricted to aspects relevant to geodetic applications that involve observations of extragalactic radio sources. In such applications, arrays of two or more radio telescopes observe any given source simultaneously. From sets of observations of a suite of such sources, one can obtain the desired geodetic information: baseline-vector, polar-motion, and earth-rotation parameters.*

We shall first describe briefly the instrumentation used in these observations and then discuss the basic observables and their simplest interpretation. Finally, we consider some complications of the interpretation due to the various geophysical "signals" and non-geophysical "noise" that affect the observables.

## Instrumentation

A VLBI system consists of an array of at least two antennas that observe the same radio source simultaneously. A direct electrical connection is not maintained between the antennas, thus allowing them to be separated by thousands of kilometers. The local-oscillator signals, used at each antenna to convert the radio-frequency signals from the source to the video (low-frequency) band, are derived from a frequency standard at the site. These standards are sufficiently stable that the relative phases of the signals from the source received at the two antennas are preserved.

*Several groups are currently engaged in such geodetic applications of VLBI: (1) A group from the Goddard Space Flight Center, the Haystack Observatory, the Massachusetts Institute of Technology, and the National Geodetic Survey; (2) a Jet Propulsion Laboratory - National Geodetic Survey group; and (3) a Canadian-British collaboration. A European consortium is also being organized for similar purposes.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

The video signals are recorded on magnetic tape at each site, with the reference time for the recordings being derived from the same standard as is used to govern the local-oscillator signals. The tape recordings are then transported to a common center where those recorded simultaneously are cross-correlated to obtain the basic VLBI observables.

## Observables

The basic observables in geodetic VLBI experiments are (i) the difference in the times of arrival at two antennas of a signal from a source; and (ii) the rate of change of this time difference. The measurement of the time-of-arrival difference can be of two types: the phase-delay difference or the group-delay difference. The phase-delay-difference observable (hereinafter "phase delay"), being based on measurements of phase, can be obtained very precisely, but usually ambiguously, due to the inability to resolve the "$2\pi n$" problem. The group-delay-difference observable (hereinafter "group delay")is usually determined with less precision than is the phase delay. But the group delay, being determined by the rate of change of phase delay with frequency, is usually unambiguous. To obtain reasonable accuracy in the measurement of group delay it is necessary to make phase-delay measurements over a wide band of frequencies simultaneously, or nearly simultaneously, since the uncertainty in the group-delay measurement is inversely proportional to this bandwidth. In practice, only a relatively narrow band of frequencies can be recorded. But this band can be split up into narrower bands which are spread over a very wide band. This technique is called bandwidth synthesis. In the newest, Mark III, system (see below), 28 narrow bands, each 2 MHz wide, are being distributed over a total band of up to 400 MHz. If the error in the measurement of phase for any one band is $\sigma(\phi)$, then the error $\sigma(\tau_g)$ in the measurement of the group delay will be given by $\sigma(\tau_g) \sim \sigma(\phi)/\Delta f$, where $\Delta f$ is the rms spread of the center frequencies of the individual bands about their mean. Because these individual bands do not cover the entire spanned band, the estimate of the group delay, too, could be ambiguous. However, a proper choice of the spacing of the individual bands, as explained below, can insure that any inherent ambiguity can be eliminated reliably.

The difference of the phases of the

signals that would be received at two antennas as a function of frequency exhibits curvature due primarily to the effects of the earth's ionosphere: the lower the frequency, the shorter the phase delay. The actual shape of the phase vs. frequency curve for VLBI observations will depend on the relative amounts of plasma between the source and the two sites. But with appropriate spacings of the narrow bands, one can "connect" unambiguously the phase at one band with those at all the other bands. For example, bands spaced in accord with a geometric series allows use of a bootstrap technique to connect phases first between the closest bands and then between the more widely separated bands through use of the characteristics of the curve established by the connection between the closest bands. Of course, the "absolute" phase would still be uncertain by multiples of $2\pi$, but the relative phases between bands would be freed from any such ambiguity. Only the relative phases affect the group delay which is equal to the slope of the curve of phase delay vs. (angular) frequency.

## Information Content

We now consider the information content of the observables under simplified assumptions. In particular, let us ignore the propagation medium and assume that the earth is rigid and rotates with a constant, known, angular velocity. We may then write the expression for the delay observable as a function of time as:

$$\tau(t) = \frac{1}{c}\vec{B}(t)\cdot\hat{s} + \tau_o^{cl} + \dot{\tau}_o^{cl}\,[t-t_o] \quad , \qquad (1)$$

where $\vec{B}$ is the baseline vector connecting a pair of antennas, $\hat{s}$ is a unit vector in the direction of the source, and $\tau_o^{cl}$ and $\dot{\tau}_o^{cl}$ are, respectively, the offsets in epoch and rate of the clock at one site with respect to those at the other site. This equation represents a diurnal sinusoid added to a straight line. The first term in the equation contributes the diurnal sinusoid due to the rotation of the baseline vector in inertial space. The slope of the straight line is due to the clock-rate offset and the intercept is due to a combination of the clock-epoch offset and the product of the polar components of $\vec{B}$ and $\hat{s}$. Clearly this curve can be specified by four parameters: the intercept and slope of the straight line, and the amplitude and phase of the sinusoid. (We do not consider the period of the sinusoid, since that is given by assumption.) Thus four measurements of the delay suffice, in principle, to determine $\tau(t)$; any additional measurements will be redundant. But how many unknown parameters are there for this situation? Naively, one would conclude that the baseline vector contributes three, the source two, and the clocks two, for a total of seven. However, the origin of right ascension of our system is arbitrary; only the origin of declination is fixed by the assumption of a known angular velocity for the earth. Since the right ascension of the source can be used to define this arbitrary origin, the number of unknown parameters is only six. Nonetheless, a unique solution cannot be obtained for these six parameters from observations of a single source. Observations of each additional source adds two unknowns: the coordinates of the source on the plane of the sky. But such observations also can be used to determine three additional parameters: the intercept of the straight line and the amplitude and phase of the sinusoid appropriate for the additional source. The slope of the straight line provides no new information since it is determined solely by the clock-rate offset. It is clear that with four observations of one source and three each of two more sources, a useful solution for all of the parameters of this simple model can in general be determined. (The measurements of delay rates simultaneously do not reduce the requirement for observations of three sources.) The accuracy of the determination of these ten parameters will depend, of course, not only on the accuracy of the measurements of delay (and delay rate), but also on the baseline, the distribution of the sources in the sky, and the distribution of the observations in time.

## Complications

The situation actually encountered with VLBI is, of course, far more complicated than outlined in the previous section. We can conveniently divide these complications into two categories: signals and noise. Here signals refer to those effects on the observables which are of geophysical interest, and noise refers to those of no intrinsic interest. (This point of view, needless to say, is a rather parochial one since one person's noise is often another's livelihood.)

We shall consider precession, nutation, solid-earth tides, crustal motions, variations in UT1, and polar motion to be signals. On the other hand, clock instabilities and uncertainties in our knowledge of source characteristics and of the propagation medium shall be considered as noise. We discuss each set in turn.

### Signals

Precession and Nutation. Changes in the direction in space of the spin axis of the earth with periods long compared to a day are sensed with VLBI through the corresponding changes in the coordinates of the radio sources. These changes will, however, preserve the arclengths between sources. At present, estimates of the precession constant made from VLBI measurements have an uncertainty of a few tenths of an arcsecond per century, severalfold larger than the uncertainty associated with the presently accepted value based on optical observations. The VLBI estimate is consistent with the optical one to within twice the formal standard error of the former. No estimates of any of the nutation terms have yet been made.

Solid-Earth Tides. The semi-diurnal solid-earth tide imparts a distinctive signature to the VLBI observable, since almost all other effects introduce a diurnal signature. The maximum magnitude of this effect on the observable has been slightly greater than one nanosecond. Thus estimates of the local values of the vertical and horizontal Love numbers, $\ell$ and $h$, can be obtained from VLBI data. Current estimates agree, to within their uncertainty of about 0.05, with the "expected" values.

Crustal Motion. Changes in crustal configuration can be sensed by long-term changes in baseline lengths; in addition, significant changes in the corresponding baseline directions for an array of antennas would signify crustal motions provided that these changes were incompatible with a rigid rotation of the array. No measurements of crustal motions have yet been obtained from VLBI data, but there is every reason to believe that such motions will be detected within a few years.

UT1 and Polar Motion. Variations in the rate of rotation of the earth and in the position of the axis of figure with respect to the axis of rotation affect the directions of baseline vectors. As with the arclengths between sources with respect to precession and nutation, the lengths of baselines are unaffected by such variations in the rate of rotation of the earth and in the position of the pole.

We should stress that the VLBI observables, for an arbitrary baseline, have no sensitivity to the "initial" orientation of the earth and direction (in space) of its axis of figure. Only changes in these quantities can be detected. In addition, any "common-mode" errors in the epoch settings of the clocks at the antenna sites will be indistinguishable from corresponding changes in the orientation of the earth about its spin axis (UT1). Finally, note that VLBI data obtained for one baseline are sensitive to only two independent combinations of the three parameters needed to specify changes in the position (in space) of the axis of figure of the earth and in the orientation of the earth about this axis: Such changes affect only the direction of the baseline which is described by only two independent parameters. Consider, as an example, a wholly north-south baseline. With such a configuration, the VLBI data would have no sensitivity to a rotation of the earth about the pole. For an east-west baseline, on the other hand, polar motion in a direction along the meridian of the midpoint of the baseline would not affect the VLBI observables. In both these cases, the baselines would undergo parallel displacements which cannot be detected from observations of sources "at infinity." Two baselines, or at least three antennas, are needed in a VLBI array to detect all three components of the changes in UT1 and pole position. These baselines must, of course, not be parallel.

Estimates of UT1 and polar motion from VLBI data now have accuracies comparable to those of other techniques, such as the classical optical methods, the Doppler tracking of satellites (for polar motion), and the laser ranging to retroflectors on the moon (for UT1). It is expected that the accuracy of the VLBI estimates will improve nearly tenfold within the next five years due primarily to the introduction and use of the new, Mark III, VLBI system. A prototype of this system has already been tested successfully; five copies of the complete system are currently under construction for placement at suitably distributed antennas. It will be important to check the improved determinations of polar motion and UT1 through redundancy and through comparison with the results from other improved techniques in order to assess the accuracy of these determinations.

## Noise

Clock Instabilities. The two parameters, for clock epoch and rate offsets, do not provide an adequate representation of the relative behavior of the clocks at any two antennas of an interferometric array over the period of many hours needed to determine the baseline vector, source positions, etc. This statement applies to the current field units of all atomic clocks, including the hydrogen-maser frequency standards.

A number of possibilities exists to minimize the impact of these clock instabilities on the accuracy with which geophysical information can be extracted from VLBI data. First, one can use higher-order polynomials to represent the relative clock behavior; here the point of "diminishing returns" sets in at about the sixth order. Second, the clock performance, especially of hydrogen-maser standards, can be improved to match that achieved in the laboratory. The sensitivity of the maser standards to environmental effects can also be reduced to minimize the introduction of diurnal signatures into the VLBI data. Third, one can reduce the effects of long-term drifts in the relative clock behavior by using clock stars. Thus, one can make observations repeatedly, say every hour, of some suitable source and use these observations to correct for the relative clock drifts. To be suitable, this source should be visible from both sites for a large fraction of the diurnal cycle and should yield a reasonably large correlated flux density so that accurate delay observations are possible to make.

Source Characteristics. The radio sources affect the determination of geophysically interesting quantities through the strength of their radio emissions, their distribution on the sky, and the accuracy with which we can determine their positions. These positions, in turn, depend on the structure and internal kinematics of the regions of radio emission in each object.

At present, the entries in the catalog of known, and potentially-usable, extragalactic radio sources number in the hundreds. Positions of a few dozen of those with the strongest emissions are now being determined routinely with an estimated accuracy of about 0".02, except for the declinations of sources that lie near the equatorial plane. The accurate determination of the declination of those sources requires the use of interferometers with baselines that possess large components in the north-south direction. Few such baselines have so far been available for extensive sets of measurements.

Aside from the examination of the characteristics of the postfit residuals, the main method for assessment of the accuracy of source-position determinations is the comparison of results obtained with different equipment. Such comparisons, made several years ago and based on data obtained with somewhat less advanced VLBI systems, showed agreement to within about 0".05 rms. (Note that a 0".001 error in source position corresponds approximately to a two-centimeter error in length for a 4,000-km baseline.)

Most extragalactic radio sources are not "points" when viewed on the scale of milliarcseconds. Rather, they exhibit complicated structure. This structure in their brightness can be mapped and a suitable feature in the map, or the overall center of brightness, can be used as a reference point. There are, however, technical difficulties in the determination of unambiguous brightness maps. These difficulties are being overcome and reliable maps on the scale of tenths of a milliarcsecond are now being obtained for some of the radio sources.

There is yet a further difficulty in the use of extragalactic radio sources: most are not static. Dramatic changes have been observed in the brightness structure of some of these sources at the level of tenths of a milliarcsecond in angular resolution on a time scale of a few months. Thus, to enable positions of extragalactic radio sources to be used effectively as a reference system at the level of milliarcsecond accuracy for geophysical applications of VLBI, one must monitor the brightness distributions of these sources as a function of time and, perhaps, as a function of radio frequency as well.

Propagation Medium. In regard to any substantial effects on VLBI data, the propagation medium can be considered to be composed of two components: the ionosphere and the troposphere. The effects of the ionosphere can, and will, be virtually eliminated by observing simultaneously in two widely separated radio frequency bands ($\simeq$2GHz and $\simeq$8GHz). The Mark III VLBI system is equipped for such dual-band observations. Moreover, enough suitable sources exist to allow effective use of the dual-band technique.

The troposphere is in effect non-dispersive at radio frequencies and is therefore a more troublesome contributor of noise. The troposphere can also be decomposed into two components: wet and dry. For the latter, the assumption of hydrostatic equilibrium is a very good one; measurements at each site of surface pressure combined with a good model of the atmosphere, then allows a good estimate to be made of the phase delay added in the zenith direction by the dry component: about 7.5 nanoseconds (equivalent to an increase in path length of about 2.3 m). It is widely thought that the error in this estimate can be kept at the 0.1% level or perhaps below. Mapping to other zenith angles, however, will increase the error somewhat since the "slant" atmospheric path length cannot be determined so accurately from the pressure measurement at the antenna site. The situation with the wet component is more difficult. The water-vapor in the atmosphere is not in hydrostatic equilibrium and is quite variable in amount. Although the total effect on the path length of radio waves is, on average, only about 7% of that of the dry component, the wet component cannot be modeled accurately. Various simple techniques have been used to try to ameliorate this problem. Such techniques involve various combinations of model atmospheres and mapping functions with or without dependence on surface measurements of temperature, pressure, and dew point, and with or without parameters that can be estimated for each site from short, ~8 hr, spans of data. Unfortunately, these techniques may well be deficient, especially for long baselines, in removing the effects of the atmosphere on the estimates of the "vertical" component of the baseline with which they are highly correlated.

The technique which has elicited the greatest expectations for providing the solution to the wet-component problem is based on the use of radiometer measurements at each site of the brightness temperature of the atmosphere at and near the ≈23GHz line in the spectrum of water-vapor emission. Studies indicate that this brightness temperature can be related with reasonably high accuracy to the excess path length attributable to the water-vapor content of the atmosphere. However, to date, almost all VLBI results have been obtained without the benefit of water-vapor radiometer measurements.

Atmospheric effects thus loom as the limiting factor in the accuracy achievable with VLBI in the determination of geophysical quantities. What will that limit be? An assessment based on theory alone is unlikely to be accurate. Measurements are clearly called for. Series of VLBI experiments should be made with supplementary water-vapor radiometer measurements, under a variety of local weather conditions, and for various baseline lengths. For short baselines, up to several kilometers in length, independent determination of the baseline vector can usually be made, with some effort, at the millimeter level of accuracy by means of conventional survey techniques. For long baselines, up to several thousand kilometers in length, independent means of verification at the relevant level of accuracy seem to be limited to laser ranging to artificial satellites or to the moon; such verification, however, will not be easy nor inexpensive for a number of practical reasons. The repeatability and consistency of VLBI results themselves may well have to provide the main standards. Since suitably accurate, and independent, estimates of UT1 and polar motion may not be available, repeated checks on the individual components of the vector baseline will likely require multi-site experiments, say with four or more separated antennas, to reduce the confusion between UT1 and pole position changes on the one hand and changes in baseline direction on the other. The many antenna sites serve to over-determine UT1 and polar motion, with the redundancy providing the meaningful check on the consistency of the estimates of some of the baseline components. For some combination of the precision and the time spanned by the sets of measurements, one must be concerned also about the genuine changes in baselines expected from plate tectonics; of course, detection of such changes are a major purpose of the measurements.

At present, checks on the repeatability of baseline determinations have involved primarily two-element interferometers. For short baselines, of the order of one kilometer in length, repeatability has been obtained at the five millimeter level in all components, and verified later by the results of a conventional survey. For long baselines, of the order of several thousand kilometers in length, only repeatability in baseline length has been meaningful; here the spread about the mean in a recent series of a dozen sets of measurements was under five centimeters. The source positions used in the analysis of each of these experiments were fixed in accord with the results from the ensemble of experiments; errors in these positions tended therefore not to have a serious effect on the repeatability of the determinations of baseline length.

## Conclusion

The future of VLBI as applied to geodetic and geophysical problems, especially to the determination of UT1 and polar motion, looks quite bright. Although the last ten years have been devoted almost exclusively to the development of VLBI, the next ten should yield significant results.

## Acknowledgment

This work was supported in part by the National Science Foundation and in part by the National Aeronautics and Space Administration.

## Bibliography

Cannon, W. H., et al., 1979, J. Geophys. Res., 84, in press.

Cohen, M. H., et al., 1977, Nature, 268, 405.

Clark, T. A., et al., 1976, Astron. J., 81, 599.

Ong, K. M., et al., 1976, J. Geophys. Res., 81, 3587.

Robertson, D. S., et al., 1979, Proc. of IAU Colloq. No. 82, in press.

Rogers, A. E. E., et al., 1978, J. Geophys. Res., 83, 325.

Shapiro, I. I., et al., 1974, Science, 186, 920.

Thomas, J. B., et al., 1976, J. Geophys. Res., 81, 995.