Chapter 17

Basic Principles of Radio Astrometry

Alain Baudry

Observatoire de Bordeaux, BP 89, F-33270 Floirac

17.1 Introduction and Basic Formalism

Modern astrometry aims at improving our knowledge of celestial body positions, motions and distances to a high accuracy. The quest for accuracy began in the early days of astronomy and is still continuing in the optical domain with most sophisticated instruments (automated meridian circles, the Hipparcos satellite or future astrometric space missions) as well as in the radio domain (connected-element interferometers and VLBI). New instrumental concepts or calibration procedures and increased sensitivity are essential to measure highly accurate positions of stars and radio sources. Positions accurate to about one thousandth to one tenth of an arcsecond have now been obtained for hundreds of radio sources and for about 100 000 to one million stars in the Hipparcos and Tycho catalogues respectively.

In this lecture we are concerned with some basic principles of position measurements made with synthesis radio telescopes and with the IRAM interferometer in particular. More details on interferometer techniques can be found in the fundamental book of [Thompson et al 1986]. The impact of VLBI in astrometry and geodesy is not discussed here. (For VLBI techniques see [Sovers et al 1998])

We first recall that measuring a position is a minimum prerequisite to the understanding of the physics of many objects. One example may be given for illustration. To valuably discuss the excitation of compact or masing molecular line sources observed in the direction of late-type stars and HII regions sub-arcsecond position measurements are required. This is because the inner layers of circumstellar envelopes around late-type stars have sizes of order one arcsecond or less and because several compact HII regions have sizes of one to a few arcseconds only. Position information is crucial to discuss not only the respective importance of radiative and collisional pumping in these line sources but also the physical asssociation with the underlying central object.

The output of an interferometer per unit bandwidth at the observing wavelength λ is proportional to the quantity

$$R = \int A(\vec{k}) I(\vec{k}) \cos(2\pi \vec{B}.\vec{k}/\lambda) d\Omega \tag{17.1}$$

where $\vec{k}$ is the unit vector toward the observed source, A is the effective antenna aperture, I the source brightness, and $\vec{B}$ the baseline vector of the interferometer. For an extended source one refers the observations to the reference direction $\vec{k}_o$ and supposing that the radiation comes from a small portion of the sky we have $\vec{k} = \vec{k}_o + \vec{\sigma}$ where $\vec{\sigma}$ is the position vector describing the source coordinates. (Since both $\vec{k}_o$ and $\vec{k}$ are unit vectors we obtain $\vec{k}_o.\vec{\sigma} = 0$.) The interferometer output is given by

$$R = V\cos(2\pi\vec{B}.\vec{k}_o/\lambda + \Psi) \tag{17.2}$$

where

$$V\exp(i\Psi) = \int A(\vec{\sigma})I(\vec{\sigma})\exp(i2\pi\vec{b}.\vec{\sigma})d\Omega \tag{17.3}$$

is the complex source visibility and $\vec{b}(u,v)$ is the baseline vector projected on a plane normal to the tracked direction. The exact definition of the baseline coordinates u and v is given in Section 17.3.

The astrometry domain corresponds to those cases where the source visibility amplitude is equal to 1 (point-like sources) and the phase provides the source position information.

17.2 The Phase Equation

The most important measurement for radio astrometry is that of the actual fringe phase of a connected-element interferometer (or similarly the group delay in VLBI). Let θ be the angle between the reference direction and the meridian plane of a given interferometer baseline. The phase is then defined by

$$\phi_r = 2\pi B\sin(\theta)/\lambda \tag{17.4}$$

If the point-like source of interest is offset by $\Delta\theta$ from the reference direction the total phase is

$$\phi = 2\pi B\sin(\theta + \Delta\theta)/\lambda \simeq \phi_r + 2\pi B\cos(\theta)\Delta\theta/\lambda \tag{17.5}$$

It is thus clear that measuring an angle or an offset position on the celestial sphere becomes possible only when all phase calibration problems have been understood and solved.

Accounting for uncertainties in the baseline and source position vectors the actual phase is

$$\phi = 2\pi(\vec{B} + \delta\vec{B}).(\vec{k}_o + \delta\vec{k})/\lambda \tag{17.6}$$

where $\vec{B}$ is a first approximation of the baseline, $\vec{k}_o$ the tracking direction; $\vec{B} + \delta\vec{B}$ and $\vec{k}_o + \delta\vec{k}$ are the true baseline and source position vectors, respectively. The reference phase is given by

$$\phi_r = 2\pi\vec{B}.\vec{k}_o/\lambda \tag{17.7}$$

and, neglecting the term involving $\delta\vec{B}.\delta\vec{k}$, we obtain

$$\phi - \phi_r = 2\pi(\vec{B}.\delta\vec{k} + \delta\vec{B}.\vec{k}_0)/\lambda \tag{17.8}$$

We consider all vector projections in the right-handed equatorial system defined by the unit vectors a_1 ($H =$ 6 h, $\delta = 0$), a_2 ($H = 0$ h, $\delta = 0$), a_3 ($\delta = 90°$). H and δ are the hour angle and declination, respectively. In this coordinate system the baseline vector $\vec{B}$ has components (B_1, B_2, B_3) and the components of the reference position $\vec{k}_o$ are given by $(\cos(\delta)\sin(H), \cos(\delta)\cos(H), \sin(\delta))$

The two limiting cases $\delta\vec{k} = 0$, and $\delta\vec{B} = 0$ correspond to those where we either calibrate the baseline or determine the exact source position.

In the first case the source coordinates are perfectly known and by comparing the observed phase ϕ with the reference phase ϕ_r one determines $\delta\vec{B}$ and hence the true baseline $\vec{B} + \delta\vec{B}$. The reference sources observed for baseline calibration are bright quasars or galactic nuclei whose absolute coordinates are accurately known. The most highly accurate source coordinates are those of the radio sources used to realize by VLBI the International Celestial Reference Frame (ICRF); distribution of coordinate errors are below one milliarcsecond. However, the ICRF catalogue is insufficient for phase and baseline calibrations of millimeter-wave arrays because most sources are not bright enough in the millimeter-wave domain. The IRAM calibration source list is thus a combination of several catalogues of compact radio sources.

17.3 Determination of Source Coordinates and Errors

Once the baseline is fully calibrated ($\delta\vec{B} = 0$) the exact source coordinates are known from the $\delta\vec{k}$ vector components. These components are formally deduced from the differential of $\vec{k}_o$

$$\begin{aligned}\delta\vec{k} &= (-\sin(\delta)\sin(H)\Delta\delta - \cos(\delta)\cos(H)\Delta\alpha, \\ &-\sin(\delta)\cos(H)\Delta\delta + \cos(\delta)\sin(H)\Delta\alpha, \\ &\cos(\delta)\Delta\delta)\end{aligned} \tag{17.9}$$

where $\Delta\alpha$ and $\Delta\delta$ are the right ascension and declination offsets in the equatorial system ($\Delta\alpha = -\Delta H$). The phase difference is then a sinusoid in H

$$\frac{(\phi - \phi_r)\lambda}{2\pi} = \vec{B}\delta\vec{k} = A\sin(H) + B\cos(H) + C \tag{17.10}$$

where

$$A = -B_1\sin(\delta)\Delta\delta + B_2\cos(\delta)\Delta\alpha \tag{17.11}$$

$$B = -B_2\sin(\delta)\Delta\delta - B_1\cos(\delta)\Delta\alpha \tag{17.12}$$

$$C = B_3\cos(\delta)\Delta\delta + \phi_{ins} \tag{17.13}$$

and C contains the instrumental phase ϕ_{ins}.

Measurement of the phase at time intervals spanning a broad hour angle interval allows us to determine the three unknowns A, B, and C, and hence $\Delta\alpha$ and $\Delta\delta$ and the exact source position. Note that for sources close to the equator, A and B alone cannot accurately give $\Delta\delta$. In the latter case, C must be determined in order to obtain $\Delta\delta$; this requires to accurately know the instrumental phase and that the baseline is not strictly oriented along the E-W direction (in which case there is no polar baseline component).

A synthesis array with several, well calibrated, baseline orientations is thus a powerful instrument to determine $\delta\vec{k}$. In practice, a least-squares analysis is used to derive the unknowns $\Delta\alpha$ and $\Delta\delta$ from the measurements of many observed phases ϕ_i (at hour angle H_i) relative to the expected phase ϕ_r. This is obtained by minimizing the quantity

$\Sigma(\Delta\phi_i' - (A\sin(H_i) + B\cos(H_i) + C))^2$ with respect to A , B, and C where $\Delta\phi_i' = (\phi_i - \phi_r)\lambda/2\pi$. A complete analysis should give the variance of the derived quantities $\Delta\alpha$ and $\Delta\delta$ as well as the correlation coefficient.

Of course we could solve for the exact source coordinates and baseline components simultaneously. However, measuring the baseline components requires to observe several quasars widely separated on the sky. At mm wavelengths where atmospheric phase noise is dominant this is best done in a rather short observing session whereas the source position measurements of often weak sources are better determined with long hour angle coverage. This is why baseline calibration is usually made in separate sessions with mm-wave connected-element arrays.

The equation giving the source coordinates can be reformulated in a more compact manner by using the components u and v of the baseline projected in a plane normal to the reference direction. With v directed toward the north and u toward the east, the phase difference is given by

$$(\phi - \phi_r) = 2\pi(u\cos(\delta)\Delta\alpha + v\Delta\delta) \tag{17.14}$$

Comparing this formulation to the sinusoidal form of the phase difference we obtain

$$u = (-B_1\cos(H) + B_2\sin(H))/\lambda \tag{17.15}$$

$$v = (B_3\cos(\delta) - \sin(\delta)(B_1\sin(H) + B_2\cos(H)))/\lambda \tag{17.16}$$

Transforming the $B_{1,2,3}$ into a system where the baseline is defined by its length $B = (B_1^2 + B_2^2 + B_3^2)^{0.5}$ and the declination d and hour angle h of the baseline vector (defined as intersecting the northern hemisphere) we obtain

$$B_1 = B\cos(d)\sin(h), B_2 = B\cos(d)\cos(h), B_3 = B\sin(d) \tag{17.17}$$

and

$$\begin{aligned} u &= (\cos(d)\sin(H-h))B/\lambda \\ v &= (\cos(\delta)\sin(d) - sin(\delta)\cos(d)\cos(H-h))B/\lambda \end{aligned} \qquad (17.18)$$

which shows that the locus of the projected baseline vector is an ellipse.

In order to derive the unknowns $\Delta\alpha$ and $\Delta\delta$ the least-squares analysis is now performed using the components u_i, v_i derived at hour angle H_i. In the interesting case where the phase noise of each phase sample is constant (this occurs when the thermal noise dominates and when the atmospheric phase noise is "frozen") one can show that the error in the coordinates takes a simple form. For a single baseline and for relatively high declination sources the position error is of order $\sigma_\varphi/(2\pi\sqrt{n_p}(B/\lambda))$ where σ_φ is the phase noise and n_p the number of individual phase measurements. This result implies (as expected a priori) that lower formal uncertainties are obtained with longer observing times and narrower synthesized beams. Of course the position measurements are improved with several independent interferometer baselines; the precision improves as the inverse of the square root of $n(n-1)/2$ for n antennas in the array.

We have shown that for a well calibrated interferometer the least-squares fit analysis of the phase in the (u, v) plane can give accurate source coordinates. However, the exact source position could also be obtained in the Fourier transform plane by searching for the coordinates of the maximum brightness temperature in the source map. The results given by this method should of course be identical to those obtained in the (u, v) plane although the sensitivity to the data noise can be different.

Finally, it is interesting to remind that the polar component of the baseline does not appear in the equation of the fringe frequency which is deduced from the time derivative of the phase. There is thus less information in the fringe frequency than in the phase.

17.4 Accurate Position Measurements with the IRAM Interferometer

Let us start with two general and simple remarks. First, the phase equation giving the angular offset θ in Sec.17.2 shows that higher position accuracy (namely smaller values of the angular offset) is achieved for smaller values of the fringe spacing λ/B. (This was demonstrated above in the case of the least squares analysis of the u, v data.) Thus, for astrometry it is desirable to use long baselines and/or to go to short wavelengths. However, the latter case implies that the phases are more difficult to calibrate especially at mm wavelengths where the atmospheric phase fluctuations increase with long baselines. Sensitivity is always important in radio astrometry. For a point-like or compact source the sensitivity of the array varies directly as $D^2(n(n-1))^{0.5}$ where D is the antenna diameter and n is the number of antennas. Thus the detection speed varies as $D^4 n(n-1)$ and big antennas are clearly advantageous. Comparison of the IRAM 5-element array with one of its competitors, OVRO with 6 x 10.4 m, gives a ratio of detection speed of 1 over 0.35 in favor of the Plateau de Bure array. (Note also that the sixth antenna in the Bure array will increase the detection speed by 50%.) In addition, the large dishes of the IRAM array are good to perform quick baseline and phase calibrations; this is another clear advantage of the IRAM interferometer in astrometric observations.

17.4.1 Absolute positions

To illustrate the potential of the IRAM array for astrometry we consider here observations of the SiO maser emission associated with evolved late-type stars. Strong maser line sources are excited in the $v = 1, J = 2-1$ transition of SiO at 86 GHz and easily observed with the sensitive IRAM array. Because of molecular energetic requirements (the vibrational state $v = 1$ lies some 2000 K above the ground-state) the SiO molecules must not be located too much above the stellar photosphere. In addition, we know that the inner layers of the shell expanding around the central star have sizes of order one arcsecond or less. Therefore, sub-arcsecond position accuracy is required to locate the SiO sources with respect to the underlying star whose apparent diameter is of order 20-50 milliarcseconds. For absolute position measurements one must primarily:

- select long baselines to synthesize small beamwidths,
- make a highly accurate baseline calibration observing several quasars selected for their small position errors,
- observe at regular intervals two or more quasars (phase calibrators) in the field of each program star in order to determine the instrumental phase and to correct for atmospheric phase fluctuations,
- observe the program star over a long hour angle interval, and use the best estimate of the stellar coordinates.

Our first accurate radio position measurements of SiO masers in stars and Orion were performed with the IRAM array in 1991/1992. We outline below some important features of these observations [Baudry et al 1994]. We used the longest E-W baseline available at that time, about 300 m, thus achieving beams of order 1.5 to 2 arcseconds. The RF and IF bandpass calibrations were made accurately using strong quasars only. To monitor the variable atmosphere above the array and to test the overall phase stability, we observed a minimum of 2 to 3 nearby phase calibrators. Prior to the source position analysis we determined accurate baseline components; for the longest baselines the r.m.s. uncertainties were in the range 0.1 to 0.3 mm. The positions were obtained from least-square fits to the imaginary part of the calibrated visibilities. (Note that the SiO sources being strong, working in the (u, v) or image planes is equivalent.)

The final position measurement accuracy must include all known sources of uncertainties. We begin with the formal errors related to the data noise. This is due to finite signal to noise ratio (depending of course on the source strength, the total observing time and the general quality of the data); poorly calibrated instrumental phases may also play a role. In our observations of 1991/1992 the formal errors were around 10 to 30 milliarcseconds. Secondly, phase errors arise in proportion with the baseline error $\delta\vec{B}$ and the offset between the unit vectors pointing toward the stellar source and the nearby phase calibrator. This phase error is $\delta(\phi-\phi_r) = (\delta\vec{B}.(\vec{k}_{quasar}-\vec{k}_o))2\pi/\lambda$. Typical values are $\delta B \simeq 0.2$ mm and $\delta k \sim 10^\circ - 20^\circ$ corresponding to phase errors of 3° to 7°, that is to say less than the typical baseline residual phases. A third type of error is introduced by the position uncertainties of the calibrators. This is not important here because the accuracy of the quasar coordinates used during the observations were at the level of one milliarcsecond.

The quadratic addition of all known or measured errors is estimated to be around $0.07''$ to $0.10''$. In fact, to be conservative in our estimate of the position accuracy we measured the positions of nearby quasars using another quasar in the stellar field as the phase calibrator. The position offsets were around $0.1''$ to $0.2''$ depending on the observed stellar fields; we adopted $0.1''$ to $0.2''$ as our final position accuracy of SiO sources. The SiO source coordinates are derived with respect to baseline vectors calibrated against distant quasars. They are thus determined in the quasi-inertial reference frame formed by these quasars.

Finally, it is interesting to remind a useful rule of thumb which one can use for astrometry-type projects with any connected-element array provided that the baselines are well calibrated and the instrumental phase is stable. The position accuracy we may expect from a radio interferometer is of the order of 1/10th of the synthesized beam (1/20th if we are optimistic). This applies to millimeter-wave arrays when the atmospheric fluctuations are well monitored and understood. With baseline lengths around 400 m the IRAM array cannot provide position uncertainties much better than about $0.05 - 0.1''$ at 86 GHz. Extensions to one kilometer would be necessary to obtain a significant progress; the absolute position measurements could then be at the level of 50 milliarcseconds which is the accuracy reached by the best optical meridian circles.

17.4.2 Relative Positions and Self-calibration Techniques

We have measured with the IRAM array the absolute position of the SiO emission sources associated with each spectral channel across the entire SiO emission profile. Any spatial structure related to the profile implies different position offsets in the direction of the star. Such a structure with total extent of about 50 milliarcseconds is observed in several late-type stars. This is confirmed by recent VLBI observations of SiO emission in a few stars. VLBI offers very high spatial resolution but poor absolute position measurements in line observations.

The best way to map the relative spatial structure of the SiO emission is to use the phase of one reference feature to map all other features. This spectral self-calibration technique is accurate because all frequency-independent terms are cancelled out. The terms related to the baseline or instrumental phase uncertainties as well as uncalibrated atmospheric effects are similar for all spectral channels and cancel out in channel to channel phase differences. By making the difference

$$(\phi(\nu) - \phi(\nu_{ref}))(\lambda/2\pi) = (\vec{B}.\delta\vec{k}(\nu) - \vec{B}.\delta\vec{k}(\nu_{ref}) \qquad (17.19)$$

where the SiO reference channel is at frequency ν_{ref} we obtain a phase difference equation whose solution gives the coordinate offsets $\Delta\alpha(\nu)$ and $\Delta\delta(\nu)$ relative to channel ν_{ref}. The main limitation in such self-calibration techniques comes from the thermal noise and the achieved signal to noise ratio SNR. The angular uncertainty $\Delta\theta$ can then be estimated with the simple equation $\Delta\theta = 0.5(\lambda/B)/\mathrm{SNR}$. Common practice with connected-element arrays shows that selection of a reference channel is not critical; it must be strong in general. Self-calibration proved to be successful with the IRAM array in several stars and in Orion where we have obtained very detailed relative maps of SiO emission. Detailed relative maps were also obtained for the rare isotope ^{29}SiO; this emission is nearly 2 orders of magnitude weaker than that of the main isotope.

The relative spot maps obtained with connected-element arrays do not give the detailed spatial extent of each individual channel. This would require a spatial resolution of about one milliarcsecond which can only be achieved with VLBI techniques. Note however that VLBI is sensitive to strong emission features while the IRAM array allows detection of very weak emission; thus the two techniques appear to be complementary.

With SiO spatial extents of about 50 milliarcseconds and absolute positions at the level of 0.1 arcsecond it is still difficult to locate the underlying star. We have thus attempted to obtain simultaneously the position of one strong SiO feature relative to the stellar photosphere and the relative positions of the SiO sources using the 1 and 3 mm receivers of the IRAM array. This new dual frequency self-calibration technique is still experimental but seems promising.