Elements of Geodetic and Astrometric
Very Long Baseline Interferometry

Axel Nothnagel

Axel Nothnagel
Institute of Geodesy and Geoinformation of the University of Bonn, Nußallee 17,
D-53115 Bonn, e-mail: nothnagel@uni-bonn.de
# Contents

4 Elements of Geodetic and Astrometric Very Long Baseline Interferometry ................................................ 1

Axel Nothnagel

1 Introduction .............................................. 4

1.1 Overview .............................................. 4

1.2 Reference frames ..................................... 6

2 Basic theory ........................................... 8

3 Technical considerations of VLBI observations ........... 9

4 Scheduling process ..................................... 11

5 Observations and calibrations .......................... 15

5.1 Radio telescopes ..................................... 15

5.2 Telescope deformations .............................. 19

5.3 Receiving system ................................... 25

5.4 System calibrations .................................. 26

6 Correlation and fringe fitting .......................... 27

6.1 General overview ..................................... 27

6.2 Making use of the cross-correlation theorem ....... 28

6.3 Efficient ways to obtain the cross correlation and cross spectrum functions .......................... 29

6.4 Fractional sample delay, fringe rotation and quantization noise ..................................... 33

6.5 Fringe fitting and bandwidth synthesis .............. 35

7 Data analysis ............................................ 38

7.1 Functional model ..................................... 38

7.2 Relativistic VLBI model delay ....................... 39

7.3 Atmospheric refraction .............................. 42

7.4 Further model contributions ........................ 44

7.5 Parameter estimation .................................. 46

8 Results .................................................... 54

References .................................................. 56
Abstract This chapter describes the theory and the individual operational steps and components needed to carry out geodetic and astrometric Very Long Baseline Interferometry (VLBI) measurements. Pairs of radio telescopes are employed to observe far distant compact radio galaxies for the determination of the differences of the arrival times at the telescopes. From multiple observations of time delays of different radio sources, geodetic parameters of interest such as telescope coordinates, Earth orientation parameters, and radio source positions are inferred. The VLBI operation’s scheme generally consists of scheduling, observing session, correlation, and data analysis.

Key words: Interferometry, astrometry, correlation, fringe fitting, radio telescopes, telescope deformations, thermal expansion of radio telescopes, VLBI data analysis

1 Introduction

1.1 Overview

The technique of Very Long Baseline Interferometry (VLBI) was invented by astronomers with its first development steps in the mid 1960s for synthesizing a telescope aperture of several thousand kilometers improving the angular resolution by orders of magnitude for imaging extragalactic objects (e.g., Matveenko et al (1965), Broten et al (1967), Moran et al (1967), Bare et al (1967)). The first applications to geodesy and astrometry followed a few years later (e.g., Cohen and Shaffer (1971), Hinteregger et al (1972), Shapiro et al...
Astrometry is an indispensable sub-discipline of VLBI which deals with the positions of astronomical objects. Pairs of radio telescopes as far apart as an Earth diameter simultaneously receive the electromagnetic radiation of extra-galactic radio sources such as quasars and other compact radio galaxies at quasi-infinite distances. Ideal candidates of these sources are point-like and show no proper motions on decadal time scales.

Due to the large distance, the emission of the radio sources arrives on Earth as plain wavefronts and cause different arrival times at the two radio telescopes forming a baseline (Fig. 1). The difference in arrival times is the delay $\tau$ which is one of the observables of geodetic and astrometric VLBI.

To determine the delays, in a first step the incoming wavefronts at each telescope are digitized, time tagged, and recorded on some appropriate medium such as magnetic disks. After transportation to a correlator facility, the second step consists of a cross correlation and fringe fitting process for each individual observation to produce the observables. From a multitude of delay observables in different baseline - source geometries, the parameters of interest such as telescope coordinates, Earth orientation parameters, or radio source positions can be inferred. With the extremely high accuracy and
reliability of these parameters, VLBI contributes to two of the three pillars of geodesy, namely geometry and Earth rotation.

Since the recorded bandwidth is a key driver of the precision of VLBI observables as described below, it has always been the primary development challenge to increase the recorded bandwidth to a technical feasible and economically acceptable level. For this reason, high precision geodetic and astrometric VLBI started in 1979 with the first observing sessions of the then newly developed Mark-3 system Clark et al (1985). With a much wider bandwidth than before, precision of centimeters or even a few millimeters were reached on baselines of several thousand kilometers.

The technique has been engineered ever further since then, making use of growing capabilities of modern electronics, other hardware components, and analysis software including geophysical models. Efforts to exploit VLBI in a structured manner led to the foundation of the International VLBI Service for Geodesy and Astrometry (IVS) Nothnagel et al (2016) in 1999. Today, the IVS coordinates the majority of geodetic and astrometric developments and observations resulting in a manifold of synergies. The active network of the IVS and associated partners such as the Long Baseline Observatory (LBO, https://www.lbo.us), earlier known as Very Long Baseline Array (VLBA), consists of about 40 radio telescopes with about ten more under construction (Fig. 2). More details of the IVS, in particular of the observing programme can be found in Schlüter and Behrend (2007); Schuh and Behrend (2012) and Nothnagel et al (2016).

1.2 Reference frames

As with other space geodetic techniques, the coordinates of the radio telescopes are referred to a conventional Cartesian terrestrial reference frame with its origin at the geocenter and the x, y, and z axes defined as a right-handed system. From the origin, the x axis points at the intersection of the zero meridian and the equator, the z axis in the direction of the Earth’s rotation axis, and the y axis is orthogonal to both of them. Due to the fact that continental drift and tectonic deformations produce movements of the lithosphere of several millimeters per year, the telescope positions have to be attributed with a reference date and respective velocity components $v_x$, $v_y$, and $v_z$ in mm/year.

For practical purposes most VLBI analysis groups maintain their own realizations from their latest global analyses. All of them are very close to the International Terrestrial Reference Frame (ITRF) in its latest version (e.g., Altamimi et al (2016) ) to which the IVS contributes routinely as described by Vennebusch et al (2007); Böckmann et al (2010); Bachmann et al (2016). On average the differences are only on the order of a few millimeters for the
coordinates and some sub-millimeter per year differences for the velocities
with very few exceptions for telescopes with known peculiarities.

![Global network of radio telescopes for geodetic and astrometric observations](image)

**Fig. 2** Global network of radio telescopes for geodetic and astrometric observations

With VLBI, compact extra-galactic nuclei such as quasi-stellar objects
(quasars) and other galaxies emitting electromagnetic radiation in the radio
frequency domain are the primary objects to be observed. The majority of
these objects are quasars and for this reason, we will generalize the naming
of the radio sources sometimes only as quasars.

For their vast distance, most of the observed objects appear quasi-point-
like and do not exhibit any proper motions. However, with increased resolu-
tion of the interferometers, time dependent morphology of the radio sources
becomes visible and the definition of the exact location of the center of emis-
sion becomes time dependent at the sub-milliarcsecond level. In addition, the
asymmetric field of radiated energy leads to the so-called source structure ef-
fect which corrupts the phase of the signal and consequently the inferred
delays. Currently, these objects are excluded whenever possible but endeavor
also exist to correct for these effects with adequate models.

The second reference frame which is directly linked to VLBI observations
is the celestial frame consisting of the positions of the compact extra-galactic
radio sources. The positions are defined as angular right ascension and decli-
nation components of an equatorial polar coordinate system with the origin
in the solar system barycenter. Right ascension (R.A.) is counted clockwise
in the equatorial plane from 0 to 24 hours with minutes and seconds of time
while declination runs in degrees, minutes, and seconds of arc from -90° at
the south celestial pole to +90° at the north celestial pole.
For quite some time the most accurate frame has been the International Celestial Reference Frame in its second realization (ICRF-2) Fey et al (2015). However, work is in progress to make use of an additional wealth of observations, which has been accumulated between March 2009 and October 2017, and generate a new ICRF, i.e., ICRF3 Jacobs et al (2014). It is expected that the axis definition is accurate to 10–20 μas and the position uncertainties have a median error of 0.15 mas as compared to 0.5 mas in ICRF2 (Jacobs, pers. comm.).

Fig. 3 Source distribution of ICRF2 sources (blue) plus new sources in ICRF3 (magenta). The black line is the ecliptic.

2 Basic theory

Quasars and other compact extra-galactic radio sources are the primary objects observed in VLBI. Originating from physical processes within these objects they emit pure random noise with mostly a rather flat amplitude spectrum in the frequency bands observable on Earth. Although the radiation is random noise, this is also called the signal of the radio source.

Compact extra-galactic objects are found as far away as the frontiers of the universe at quasi-infinite distances. For this reason, the emission of these radio sources arrives on Earth as plain wavefronts and causes different arrival times at the two radio telescopes forming a baseline (Fig. 1). The difference in arrival times \( t_1 \) and \( t_2 \) is the delay, which is one of the observables of geodetic and astrometric VLBI. The principal connection between the time delay \( \tau \), the baseline \( \mathbf{b} \) and the unit vector in source direction \( \mathbf{k} \) is
\[ \tau = t_2 - t_1 = -\frac{1}{c} b \cdot k \]  
(1)

with \( c \) being the speed of light.

On the side of the observable \( \tau \), this concept more or less assumes a stationary geometry and the observations appear like snapshots of this situation. However, as we will see later, VLBI needs some finite integration time and the Earth of course rotates during that time.

In a slightly different conceptual approach, the radiation received by the two telescopes is considered as two identical monochromatic waves with the phase difference of the two signals rotating due to Earth rotation. This interferometer phase is another observable of the VLBI technique. While the Earth rotates, this phase difference rotates through many cycles depending on baseline length and orientation as well as on the wavelength of the observing frequency. The determination of this phase needs some integration time as well applying concepts of interferometry.

The interferometer phase is determined in the correlation and the subsequent fringe fitting process to search for the interferometer phase at a certain epoch (see Sec. 6). In a rough concept of the correlation process, the digitized signal streams of the two telescopes are cross-multiplied at various trial delays for finding a correlation maximum and the respective interferometer phase.

3 Technical considerations of VLBI observations

Turning to the observations, it is rather clear that the quasars and other compact radio sources do not emit just monochromatic radiation but a broad spectrum of frequencies. This is recorded with a bandpass of limited bandwidth and the time delay will turn into a group delay, which obeys the rule of

\[ \tau = \frac{d\phi(\nu)}{d\nu} \]  
(2)

with \( \phi \) being the phase and \( \nu \) being discrete observing frequencies. Its first derivative w.r.t. time is the delay rate

\[ \dot{\tau} = \frac{d\tau}{dt}. \]  
(3)

Together with some other parameters, the total spanned bandwidth determines the theoretical accuracy with which the group delay can be determined Whitney (1974):

\[ \sigma_\tau = \frac{1}{2\pi \cdot SNR \cdot \Delta\nu_{RMS}} \]  
(4)
The effective bandwidth of the receiving system depends on the number of channels $n$, the individual channel frequencies $\nu_i$ and the mean frequency $\nu_M$ (Thompson et al, 2007)

$$\Delta \nu_{RMS} = \sqrt{\frac{\sum_{i=1}^{n} (\nu_i - \nu_M)^2}{n}}.$$  \hspace{1cm} (5)

The signal to noise ratio (SNR) of an interferometer is the inverse of the phase noise (standard deviation) and can be calculated theoretically as

$$SNR = \frac{\eta S}{2k} \sqrt{\frac{A_1 \cdot A_2}{T_{sys1} \cdot T_{sys2}}} \sqrt{2\Delta \nu T}$$ \hspace{1cm} (6)

with $\eta = \text{digitizing loss factor} \ (0.5 - 0.7)$, $S = \text{correlated flux density of the radio source}$, $k = \text{Boltzmann’s constant} \ (1.38 \times 10^{-23} \ \text{Ws/K})$, $A_i = \text{effective antenna areas of telescope 1 and 2}$, $T_{sys} = \text{noise temperatures of the receiving systems}$, $\Delta \nu = \text{total bandwidth of the receiving system}$, $T = \text{coherent integration time}$. The actual SNR is computed during correlation (Sec. 6.3).

The efficiency of a radio telescope for VLBI observations is regularly determined and expressed as System Equivalent Flux Density (SEFD), $S_{sys}$. The SEFD represents the flux density of a fictitious radio source which would double the output power of the complete receiving system. The spectral flux density of a radio source as well as for a receiving system is given in Jansky [Jy] where $1 \text{ Jy} = 1 \times 10^{-26} \ \text{W/(Hz} \cdot \text{m}^2)$. $S_{sys}$ is computed according to Thompson et al (2007) with

$$S_{sys} = \frac{2k}{A_{eff}} T_{sys}$$ \hspace{1cm} (7)

where $A_{eff} = \text{effective antenna aperture}$, $T_{sys} = \text{system temperature}$, which is composed of the sky brightness temperature $T_{sky}$ and the noise temperature of the instrument $T_{inst}$,

$$T_{sys} = T_{sky} + T_{inst}$$ \hspace{1cm} (8)

with $T_{sky} = T_{so} \cdot \lambda^{2.55}$ and $T_{so}$ being the galactic background radiation of $60 \text{ K} \pm 20 \text{ K}$. $\lambda$ is the wavelength of the radiation.

The SEFD has the advantage that it can be measured easily at each telescope independently by determining the fractional increase in power obtained when pointing on and off a source of known flux density. SEFDs range from 60 Jy to 400 Jy for large antennas, 800 Jy to 1000 Jy for the 20 m class systems to more than 15000 Jy for telescopes of only a few meters of diameter. These values also strongly depend on the receiver noise mitigation by cryogenic systems.

With most of the parameters mentioned above being predefined, the integration time and the bandwidths are the parameters which still allow some
scope for increase and, thus, reduction in standard deviation of the delay observables. However, the integration time should be kept short. An initial reason was that the frequency standards allowed only for a limited coherence time but this has been overcome by the use of hydrogen maser clocks (H-Maser). Today, short integration times permit collecting observations at as many different spatial directions as possible for a good sampling of the atmospheres above the radio telescopes Petrachenko et al (2009).

Another parameter, which can be optimized, is the observed bandwidth, which appears as the effective bandwidth in Eq. 4 and as total bandwidth in Eq. 6. Since recording capacity has always been a limiting factor, bandwidth synthesis has been invented Rogers (1970). This technique works on the basis of extracting and processing comparatively small frequency bands (today up to 16 MHz) out of a spectrum of, e.g., 720 MHz to span the total bandwidth of the latter but reducing the need for storage and transportation to only n times 16 MHz (Fig. 4). While the two extreme frequency bands help to increase the effective bandwidth $\Delta \nu_{RMS}$ according to Eq. 5, the interposed bands support this as well but also increase the total bandwidth $\Delta \nu$ in Eq. 6 and mitigate sidelobes in the delay resolution function (Sec. 6). The sequence of frequencies is selected initially to produce from any pair of channels as many different frequency differences as possible. As shown later, the individual channels produce so-called single-band (group) delays while the combination of all channels in either S or X band yield multi-band (group) delays.

![Fig. 4](attachment:channel_allocation.png)  
**Fig. 4** Generic channel allocation for bandwidth synthesis.

### 4 Scheduling process

Observing VLBI sessions requires an active control process because it has to be guaranteed that the telescopes forming one or several baselines point at the same quasar for the same period of time. In addition, recording of the data has to be synchronized as well since registration is not continuous to make optimal use of the recording media. Under these premises, observing schedules are prepared a few weeks before the observing date. They contain
the start and stop times for every telescope in the network for every individual
observation. Assuming a network of up to 20 telescopes, multiple configura-
tions of subsets of these telescopes form so-called scans of one quasar at a
time. The composition of these subsets always depends on the visibility of
the quasar from a certain area on the Earth’s surface. In the end, a schedule
for an observing session of 24 hours may consist of thousands of scans. The
word scan is actually used if two or more telescopes simultaneously observe
the same radio source. Each scan produces $n \cdot (n - 1)/2$ individual delay
observations for $n$ telescope in the same scan.

The preparation of the observing schedules is of great importance to the
overall results because it defines the geometric configuration of the parameter
estimation. In this respect it is comparable to geodetic network optimization
(e.g., Grafarend and Sansó (1984), Amiri-Simkooei et al (2012)).

The planning of VLBI observations, in fact, is rather complicated because
there are many parameters which have to be taken into account. The first
question to answer always is whether a quasar is actually above the horizon
at the telescope. This can be deduced from a pure geometric consideration
resulting from the position of the quasar in a sky-fixed celestial reference
frame (see Sec. 1.2) and the coordinates of the telescope in the terrestrial
frame transformed into its instantaneous location in an geocentric celestial
reference frame. To first order this is just a simple rotation about the Earth
rotation axis with the local hour angle $h_{loc}$. With some generalization it
is computed with the right ascension $\alpha$, the longitude of the telescope $\lambda$,
Universal Time $UT$, and Greenwich Mean Sidereal Time (GMST) at 0h UT
according to

$$h_{loc} = \text{GMST}_{0h} + UT \cdot 1.00274 + \lambda - \alpha. \quad (9)$$

For any radio source with its right ascension $\alpha$ and declination $\delta$ and any
telescope with its geographic latitude $\Phi$ and its instantaneous hour angle the
respective azimuth $A$ and elevation $\varepsilon$ can be computed according to Mueller
(1969)

$$\tan A = \frac{-\cos \delta \sin h_{loc}}{\sin \delta \cos \Phi - \cos \delta \cos h_{loc} \sin \Phi} \quad (10)$$

$$\sin \varepsilon = \sin \Phi \sin \delta + \cos \Phi \cos \delta \cos h_{loc}. \quad (11)$$

Querying the local horizon mask of each telescope, which may be limiting the
observable section of the sky, the question is answered whether a radio source
is observable at a certain instant of time. By default, VLBI observations are
carried out down to the threshold of $\varepsilon \geq 3^\circ$.

Due to the fact that cable links between the fixed ground part and the
turning elements need to be secured from over-twist, all azimuth-elevation
telescopes are limited in their azimuth rotations. Normally, a full circle is
augmented by some reasonable arc length in clockwise and counter-clockwise
directions to allow for continuous tracking at all directions (Fig. 5). Special
care has to be taken in the scheduling process that the duration of the slew
from one azimuth to another is computed correctly taking into account that sometimes the shortest path is blocked by the antenna limits.

![Diagram of cable wrap limitations.](image)

**Fig. 5** Example of cable wrap limitations (-270° to +270°). N.B.: Position Az may be reached by two different azimuth values

For each source the duration of the observation has to be calculated. This mainly depends on a predefined minimal signal-to-noise ratio (SNR) which is normally set at a level of 25 to 35. For each source, the necessary integration time is computed according to Eq. 6 depending on the correlated flux density of the source and the sensitivity of the telescopes represented in system equivalent flux densities (SEFD) according to Eq. 7. For the damping caused by the atmosphere the SEFD at zenith ($S_{\text{zenith}}$) has to be scaled according to Gipson (2016) p. 170 by

$$S_{\text{sys}}(\varepsilon) = S_{\text{zenith}} \cdot f(\varepsilon).$$  \hspace{1cm} (12)

In its simplest form, the scaling factor $f(\varepsilon)$ is a mapping function just as $1/\sin \varepsilon$. A more complex model

$$f(\varepsilon) = \sum_{i=1}^{n} \frac{c_i}{(\sin \varepsilon)^y}.$$  \hspace{1cm} (13)

is being used in the SKED scheduling program Gipson (2016), where n is the number of terms (usually 1 - 3), $c_i$ is the coefficient for the $i$th term, and $y$ is the power of the $\sin \varepsilon$ term ($0 \leq y \leq 1$). Since it takes quite some effort to determine the coefficients and power law, these are available only for a few radio telescopes such as those of the Very Long Baseline Array (VLBA) Petrov et al (2009) with $c_1$ close to 1.

These mapping functions lead to an increase of the SEFDs already by a factor of 2 at $\varepsilon = 30^\circ$ increasing sharply below. For this reason, observations close to the horizon need much longer integration times than those near zenith.
The final observing schedule is constructed in a sequential forward mode where new observations are being selected to fulfill an optimization criterion. Most operational observing schedules are generated to optimize the sky coverage at each telescope in a way that the hemisphere above each telescope is filled homogeneously with observations as much as possible for a good estimation of the atmospheric refraction parameters (Fig. 6 and 7).

![Fig. 6](image_url) Local sky coverage of observations at telescope HART15 (South Africa)

![Fig. 7](image_url) Local sky coverage of observations at telescope NYALES20 (Svalbard, Norway)

It should be noted here that due to the fact that the atmospheric refraction parameters are often estimated as linear splines with interval lengths of, e.g., 1 hr, the sky coverage is actually evaluated only for the hour preceding the current instant.

The generation of the observing schedules is a contemporary field of research and many innovative optimization schemes have been devised e.g.,
In the scheduling process itself, we assume a start scenario where all telescopes point at the same quasar. The question to answer is which is the next observation fulfilling the optimization criterion. To answer this, for all possible candidates the next possible start time is computed from the slew rates of the telescopes and the minimum scan lengths for each baseline according to the SNR threshold. The final selection is then taken in agreement with some predefined minor options such as maximum scan length, minimum separation between observations of the same source. For more details of the geodetic scheduling procedure see Gipson (2016).

5 Observations and calibrations

5.1 Radio telescopes

Radio telescopes for geodetic and astrometric applications are basically the same as those for pure astronomical observations. The only differences are found in the receiver and backend technology. The dominant part of a radio telescope is the main reflector, often combined with a secondary reflector or sub-reflector, which serves to combine the energy of the incoming wave front in a single focal point. The construction of a radio telescope aims at locating the phase center of a so-called feed horn exactly at the focal point to guide the energy into the receiver which predominantly consists of the first stage of amplification in a cooled (cryogenic) environment.

A radio telescope is said to have a prime focus system if the radio frequency feed horn is placed at the focal point of the main reflector. Most radio telescopes, however, use a secondary focus near the vertex of the paraboloid for easier access to the receiver equipment. For this purpose either a hyperbolic (Cassegrain system, Fig. 8) or an ellipsoidal sub-reflector (Gregorian system, Fig. 9) are mounted in front of or behind the primary focal point, respectively, to concentrate the energy in a secondary focal point. The feed horn is then placed here.

The reflector optics is mounted on a pair of axes which are perpendicular to each other to reach all positions on the sky. The most common mount is the azimuth-elevation, also called alt-azimuth, axis system (Fig. 10) Baars (2007). Other systems are polar mounts, where the primary axis is in a position parallel to the Earth’s rotation axis (Fig. 11), and the X/Y mount, where the primary axis just lies in a horizontal plane (Fig. 12).

It should be mentioned here that azimuth-elevation telescopes are constructed in two different ways. The one group is a so-called wheel and track type where the whole structure moves on a circular track with at least four groups of wheels distributing the weight evenly (Fig. 13). Today, this con-
Fig. 8 Cassgrain radio telescope optics with parabolic main reflector and hyperbolic sub-reflector (dashed lines are discrete ray paths)

Fig. 9 Gregorian radio telescope optics with parabolic main reflector and ellipsoidal sub-reflector (dashed lines are discrete ray paths)

struction is mainly used for bigger telescopes of diameters larger than 20 m. As with telescopes with polar and X/Y mounts these are generally made of steel entirely. The second group consists of so-called turning-head telescopes where only the top part rotates while a tower made of concrete supports the moving parts (Fig. 14).

In geodetic and astrometric VLBI, the Earth-fixed coordinates of a radio telescope are always related to the VLBI reference point. This is a point within the structure of a telescope which is invariant to any rotations of the telescope Sovers et al (1998). To first order this is the intersection of the primary and the secondary rotation axis of the telescope. However, since these intersect only in very seldom cases, the reference point is the projection point of the secondary axis onto the primary axis (Fig. 10, 11, 12)
the two axes is called the axis offset ($AO$). It can range from a few millimeters in cases, where intersecting axes were planned but not exactly realized for constructional reasons, to several meters where required for technical reasons, in particular for polar and X/Y mounts. In seldom cases, the axis offset may even be negative resulting from the fact that the elevation axis lies a small distance behind the azimuth axis Nothnagel (2009).
Considering the travel time of the incoming wave fronts, the distance between the feed horn and the secondary axis is considered invariant to first order. It should be mentioned here already that any constant part of the signal path produces a constant time delay contribution which is treated as relative clock offset in the analysis (see Sec. 7). The axis offset, however, produces a time delay $\Delta\tau_{AO}$ which varies with elevation. In general, it depends on the unit vector in source direction $\mathbf{s}$ and the unit vector in the direction of the fixed axis $\mathbf{f}$ Nothnagel (2009):

$$\Delta\tau_{AO} = \frac{1}{c} AO \cdot \sqrt{1 - (\mathbf{s} \cdot \mathbf{f})^2}. \quad (14)$$
For azimuth-elevation telescopes the axis offset produces an extra time delay contribution of
\[ \Delta \tau_{AO} = \frac{1}{c} \cdot AO \cdot \cos \varepsilon. \] (15)
with \( \varepsilon \) being the elevation angle of the pointing, while for polar mounts the delay contribution depends on the declination of the radio source \( \delta \):
\[ \Delta \tau_{AO} = \frac{1}{c} \cdot AO \cdot \cos \delta. \] (16)
For X/Y mounts one has to distinguish in which direction the primary axis is oriented. For those in north-south direction, the delay contribution is
\[ \Delta \tau_{AO} = \frac{1}{c} \cdot AO \cdot \sqrt{1 - (\cos \varepsilon \cdot \cos \alpha)^2} \] (17)
with \( \alpha \) being the azimuth of the radio source, while for those in the east-west direction it is
\[ \Delta \tau_{AO} = \frac{1}{c} \cdot AO \cdot \sqrt{1 - (\cos \varepsilon \cdot \sin \alpha)^2}. \] (18)

5.2 Telescope deformations

It is needless to state that the radio telescopes used for geodetic and astrometric VLBI need to be stable enough in all their components for results of highest accuracy. This applies not only for the reference point itself but also for the reflecting optics. The stability of the reference point of VLBI telescopes has been studied only occasionally L"osler et al (2013, 2016). The results indicate that the telescope studied is stable close to the detection
limit of sub-millimeter accuracy. It is worth to note that all these studies are
made for turning-head telescopes where fixed parts of the construction can
be used to mount geodetic targets. For wheel and track telescopes any effect
of instability, manifested as tumbling of the reference point, cannot be sepa-
rated from the determination of the axis offset Holst et al (2018). However,
if a displacement of the reference point is actually a consequence of a tilting
of the telescope, this can be detected by identifying trends in its pointing
model. Furthermore, a tilting process may be interpreted as a spurious tele-
scope motion in the global frame which may be identified by comparison with
velocity vectors from observations of nearby permanent GNSS installations
as demonstrated on the PIETOWN radio telescope Petrov et al (2009).

In the construction process of a radio telescope, special care is taken that
the weight of the primary reflector, the quadrupod holding the sub-reflector,
and the sub-reflector are balanced by some counter-weight. The reason is
that the motors and gears should be relinquished from any torques caused
by imbalances for protection. For this reason, there is almost no shift in load
enforced on the elevation axis, and thus on the VLBI reference point, when
the telescope is tilted to different elevation angles. The only instability effect
of the VLBI reference point is then thermal expansion. This is mostly modeled
by applying the ambient temperature to the telescope dimensions Nothnagel
(2009) but examples exist where thermal expansion is actually measured on
the telescopes directly with the help of invar rods Johansson et al (1996);
Zernecke (1999).

In most cases, however, the effects of thermal expansion are being mod-
eled applying the expansion coefficients to the dimensions of the telescopes.
If we first consider only the support structure up to the elevation axis of an
azimuth-elevation telescope, this normally consists of some concrete founda-
tion and a tower produced of steel. Depending on the material, these have
two slightly different expansion coefficients. Examples are $\gamma_f$ of $1.0 \times 10^{-5}$
[1/°C] for a concrete base and $\gamma_p = 1.2 \times 10^{-5}$ for the steel tower. Applying
a conventional reference temperature of $T_0$ and a time lag of $t - \Delta t$ Noth-
nagel (2009), the elevation dependent effect on the delay of telescope $i$ can
be computed according to

$$\Delta \tau_{\text{therm},i} = \frac{1}{c} \left[ \gamma_f \cdot \left(T(t - \Delta t_f) - T_0\right) \cdot \left(h_f \cdot \sin \varepsilon\right) + \gamma_a \cdot \left(T(t - \Delta t_a) - T_0\right) \cdot \left(h_p \cdot \sin \varepsilon\right) \right].$$ (19)

This is by far the biggest effect of thermal expansion. For a telescope of
20 m diameter, the height of the elevation axis is roughly 12 m causing a
height variation effect of almost 3 mm originating from a 20° C temperature
difference between winter and summer.

Another part of the telescope, where stability counts, is the superstruc-
ture of the telescope being responsible for the path length of the signal. In
an ideal situation the total path length through the reflecting optics is con-
considered to be stable for any direction on the sky. However, thermal expansion and gravitational deformations cause the path length to change. Thermal expansion primarily leads to a change of the focal length of the telescope Artz et al (2014). In addition, the expansion of the legs of the quadrupod holding the sub-reflector or the feed horn in primary focus produces an extra path length at higher temperatures. Besides this pure geometric effect, the focal characteristics may be disturbed slightly leading to phase noise from different sections of the reflecting surfaces but these are estimated to appear only in the SNR budget and have to be neglected. Variations of the path length have adverse effects only within the time frame of a single session of 24 h duration because the mean effect is compensated for by the clock offset parameter in the estimation process (Sec. 7). With higher resolutions of the clock parameters down to 1 h and below, the effect is mitigated even more and the level of remaining path length variations is generally less than 1 mm. They can be modeled if all dimensions and material properties are known Nothnagel (2009).

The gravitational deformations of the paraboloid and the sub-reflector create different path lengths in general but also depending on the radial distance from the optical axis. They are more critical than those of thermal origin because they are purely elevation dependent and thus change from observation to observation. As has already been investigated by Clark and Thomsen (1988), for a prime focus telescope, the deformations can be separated in three different components which are all elevation angle \( \varepsilon \) dependent: a) the change of focal length \( \Delta F \), b) the movement of the vertex of the paraboloid \( \Delta V \), and c) the shift of the feed horn or the sub-reflector in radial direction \( \Delta R \) (Fig. 15). The extra path length \( \Delta L_{\text{grav}} \) can then be composed of

\[
\Delta L(\varepsilon) = \alpha_F \Delta F(\varepsilon) + \alpha_V \Delta V(\varepsilon) + \gamma \alpha_R \Delta R(\varepsilon).
\]  

The coefficient \( \gamma \) is equal to 1 for prime focus telescopes and 2 for secondary focus telescopes. The net effect of the elevation-dependent displacements \( \Delta F, \Delta V, \Delta R \) on the extra path length depends on scaling coefficients \( \alpha_F, \alpha_V, \alpha_R \). Here, the scaling coefficients \( \alpha_F \) and \( \alpha_V \) according to Clark and Thomsen (1988) are linearly dependent on \( \alpha_R \), i.e.,

\[
\alpha'_V = -1 - \alpha'_R \tag{21}
\]

\[
\alpha'_F = 1 - \alpha'_R \tag{22}
\]

where the \( V', R', F' \) indicate that these are parameters for prime focus telescopes. For Cassegrain systems as a type of secondary focus telescopes, Abbondanza and Sarti (2010) developed the relationships

\[
\alpha''_V = -1 - 2\alpha''_R \tag{23}
\]
Fig. 15 Gravitational deformations of telescope optics: Solid lines represent geometry at 90° elevation angle, dashed lines that at 0°. ∆V is the shift of the vertex of the telescope, ∆F is the change in focal length, ∆R is a shift of the sub-reflector (in case of a secondary focus telescope) or of the feed horn (in case of a primary focus telescope).

\[ \alpha''_F = 2 - 2\alpha''_R. \]  

The initial concept for the computation of \( \alpha_R \) by Clark and Thomsen (1988) assumed that all of the reflecting area of the main reflector contributes to the total change of path length in the same way. However, feed horns are always constructed in a way that they have only a limited aperture angle matching the dimensions of the reflector system to avoid unwanted spill-over from beyond the edge. Since this cannot be realized purely binary, i.e., homogeneous for the area of the reflector system and zero beyond, the feed horns are constructed with an edge taper. This realizes a roughly exponential drop of sensitivity towards the aperture angle needed to illuminate the reflector system (Fig. 16).

The function, which describes the drop in sensitivity depending on the aperture angle \( \theta \) (Fig. 15), is called the illumination function and in essence serves as a weighting function for the aperture of the telescope depending on the radial distance from optical axis, i.e., signal paths near the optical axis have a much higher weight than those near the rim of the main reflector.

The determination of the exact illumination function heavily depends on measurements of the gain at multiple aperture angles of the feed horn which can be easily transformed into metric separations of the point of incidence on the main reflector from the optical axis. Unfortunately, for most telescopes only the relation between the gain on-axis and at the edge is available. With only two points of the function, this leaves room for interpretation in the sense that the shape of the function is at the discretion of the analyst with Gaussian, binomial and cosine-squared models at hand Abbondanza and Sarti (2010); Artz et al (2014). Since this is a fairly new area of research, any further studies on this need to request more detailed measurements of the illumination functions on site.
The coefficient $\alpha'_R$ for a prime focus telescope, which scales the shift of the feed horn in primary focus Clark and Thomsen (1988), is

$$\alpha'_R = 8\pi f^2 \int_{t_1}^{t_2} I_n(t) \frac{1-t^2}{1+t^2} t \, dt$$  \hspace{1cm} (25)$$

where $t_1 = r_1/2F$ and $t_2 = r_2/2F$ with the radial distances from the vertex $r_1$ and $r_2$, measured in the aperture plane. The distance $r_1$ generally is non-zero because most radio telescopes have some sort of feed horn housing at the vertex which needs to be excluded. $I_n$ is the normalized illumination function.

For a secondary focus Cassegrain system, Abbondanza and Sarti (2010) derived

$$\alpha''_R = 2\pi (F^2 - a^2)^2 \int_{t_1}^{t_2} I_n(t) \frac{t}{t^2 F^2 - a^2} t \, dt$$  \hspace{1cm} (26)$$

with $t = r/2F$ for the scaling coefficient, again applying a weighting through the normalized illumination function $I_n$. The parameter $a$ is the reference semi-major axis of the ellipsoid at $90^\circ$ elevation.

For a Gregorian system Artz et al (2014) adopted a similar derivation but with an implicit illumination function on the basis of a cosine-squared model. Here the sub-reflector is a triaxial ellipsoid with its longest extension in the $z$ direction and with the first two semi-major axes ($a$ and $b$) being identical. If the sub-reflector is displaced, the travel time increases by a factor of

$$\alpha''_R = \int_{r_1}^{r_2} k \cdot 10^{\frac{r_2}{10r_1} \cos^2 \theta} \cdot \frac{1}{2}(D'_2 + D_3 - 2a) \, dr$$  \hspace{1cm} (27)$$
with the normalization coefficient $k$ Abbondanza and Sarti (2010). $D_2'$ the
distance between the focus of the sub-reflector and point of incidence on the
sub-reflector and $D_3$ the distance between the point of incidence and the
focal plane of the secondary focus (Fig. 17). The aperture angle $\theta$ can easily
be deduced from $r$. $c_0$ and $c_1$ are the coefficients of the $\cos^2$ function.

![Fig. 17 Components of path length in telescope optics.](image)

Every telescope has its own set of coefficients $\alpha_{F/V/R}$. A rough number
for $\alpha R''$ for a secondary focus telescope is 0.9 (1.06 for Effelsberg Artz et al
(2014) and 0.8 for Medicina or Noto Abbondanza and Sarti (2010)) which,
according to Eq. 23 and Eq. 24, lead to $\alpha V'' = -2.8$ and $\alpha F'' = 0.2$. This
indicates that changes in the position of the vertex of the paraboloid relative
to the VLBI reference point have about a ten times larger effect on the signal
path length than changes in the focal length and 3 times larger an effect than
of changes of the sub-reflector position.

Similar relations also apply for primary focus telescopes where Abbon-
danza and Sarti (2010) quote an $\alpha R'$ of about 0.7 resulting in an $\alpha V' = -1.7$
and an $\alpha R' = 0.3$ according to Eq. 21 and 22. Here, the multiplication factor
of a shift of the vertex of the paraboloid is about a factor of 2.5 larger than
that of the feed horn $\alpha R'$.

Although variations of $\Delta R$ may originate from several phenomena, we only
have to consider line-of-sight shifts of the sub-reflector or feed horn (Fig.
15) in the telescope-fixed reference frame. Position variations predominantly
happen due to gravitation when the telescope is tilted, e.g., by bending of
the struts holding the receiver Sarti et al (2009). These need to be measured
locally. Presently, the most suitable technique for measuring the deformations
of the primary reflector resulting in a $\Delta F$ is terrestrial laser scanning (TLS).
Performing these at discrete elevation angles between zenith and horizon
provide reliable estimates of the changes in focal length Sarti et al (2009);
Artz et al (2014); Holst et al (2015). In some cases, the other parameters ($\Delta R$,
$\Delta V$) can be deduced from the TLS results Artz et al (2014) but often they
need other sensors mounted on the structure of the telescope (Bergstrand, pers. comm.).

5.3 Receiving system

Today, geodetic VLBI systems mostly use concentric feed horns for two bands in the radio frequency (RF) domain, X band as primary observing frequency and S band for calibration of ionospheric refraction. From the injection into the feed horn, the quasar signal is lead into the first amplifier stage consisting of low-noise amplifiers, e.g., field-effect transistors (FET), and filtered to useful bandwidths of 500 to 1000 MHz (so-called intermediate frequencies, IF). This is done in separate chains for each frequency band or sub-band. It is then heterodyned with some local oscillator frequency, e.g., 2020 MHz for the S band signal and 8080 MHz for X band (Fig. 1). After reaching the control building, channels of separate frequencies with widths of 2, 4, 8, or 16 MHz are extracted and further mixed down to baseband for bandwidth synthesis.

For the description of a frequency band, in most cases, one of the band limiting frequencies is given together with the indicator upper sideband (USB) or lower sideband (LSB). Throughout the VLBI community, the channel reference frequencies 2225.99, 2245.99, 2265.99, 2295.99, 2345.99, 2365.99 MHz of S band and 8212.99, 8252.99, 8352.99, 8512.99, 8732.99, 8852.99, 8912.99, 8932.99 MHz of X band, all with USB are currently allocated. At X band, two more channels are often used at 8212.99 MHz (LSB) and 8932.99 MHz (LSB) to fill the historically provided channels no. 15 and no. 16. The reasons for selecting and processing individual channels are given in Sec. 2.

Digitization of the individual channels is done according to the Nyquist theorem with two samples of one bit each per Hz. "0" represents a negative voltage and "1" a voltage of zero or larger. This is called 1-bit sampling. In some setups, a second bit is added to each sample indicating whether the voltage is below or above a (half-power) voltage threshold (2-bit sampling).

The frequency allocation is subject to extension in the intermediate future because the VGOS concept (VLBI Global Observing System, Petrachenko et al (2009)) foresees that the full band between 2 and 14 GHz be observable simultaneously employing a single feed horn. The exact allocation of four individual frequency channels of 0.5 or 1 GHz bandwidth will be decided depending on recording capabilities and radio frequency interference (RFI) through strong artificial emitters affecting the receivers. The general signal processing chain beyond the receiver is very similar to the legacy system and backwards compatibility with the current S/X configuration is planned for mixed-mode operations with legacy installations.

The local oscillators for the down conversions are driven by frequency standards which normally are hydrogen masers for guaranteeing sufficient fre-
quency stability Nothnagel et al (2018). The bit streams are then time-tagged
in a so-called formatter unit included in the disk unit used for recording.

In many cases, the recording units themselves consist of a chassis, which is
normally mounted in some electronics rack, and exchangeable, transportable
disk units, each containing eight commodity disks for storage of the data. The
latter are used for intermediate storage and later transmission of the data by
Internet lines or for shipment to the correlator by courier service. These are
known as Mark5 recorders with Mark6 developments underway. Alternatively,
also RAID storage systems are in use (so-called Flexbuff). Naturally, they are
mainly used for transportation of the data to the correlators via the Internet.

Depending on the number of individual channels and the channel band-
widths, the total data volume for legacy systems may be up to 1 GigaBit per
second (Gbps) or 4.3 TerraByte (TB) per telescope at a 40% recording cycle
(the rest is used for slewing of the telescopes). For VGOS systems, with four
channels of 1 GHz bandwidth each, two polarizations and a 50% recording
cycle, the total volume per telescope is approximately 86 TB.

5.4 System calibrations

From the feed horn of the telescope, the signals pass through various com-
ponents of the receiving system before they are digitized and time tagged
(see Sec. 5.3). It is quite natural that the signals suffer from system delays
which are neither constant nor predictable Clark et al (1985). To calibrate
the system delay, a comb of phase calibration tones of 1 MHz separation is
injected near the feed horn (Fig. 18). This calibration signal is mixed with
the noise from the quasars, travels through the same electronic components,
and is finally recorded as part of the combined signal. The phase calibration
tones are extracted again as part of the correlation process (Sec. 6). The
voltage level of the phase calibration tones should be kept at about 1 to 10%
of the quasar noise to avoid disturbances of the latter.

Phase calibration serves two different purposes. The first one is that in-
herently the phase calibration marks the time epoch when a certain signal
component has actually arrived at the feed horn. This can be considered as a
timing reference. The second purpose is that for bandwidth synthesis, multi-
ple channels are processed in the system, i.e., digitized and down-converted,
and that dispersive effects occur to the phase of each channel. These lead to
spurious delays according to Eq. 2 which are variable and can be corrected
for in the fringe fitting process (Sec. 6.5).

The unit producing the phase calibration signal is located very close to
the feed horn (Fig. 18). It receives its reference frequency from the H-Maser
through a 5/10 MHz distributor which also supplies the local oscillators with
phase coherent frequencies. Since this link is subject to changes in the elec-
trical path caused by temperature variations and twisting of cables, it is
monitored by a two-way system, the so-called cable calibration, a.k.a. delay calibration, system. In fact, the 5 MHz signal bounced by the cable calibration antenna unit is modulated by a 5 kHz signal which, at the control room, is compared with a 5 kHz reference signal. The time equivalent of the phase shift is constantly monitored with a start-stop counter and recorded in the station log for calibration in the data analysis process (Sec. 7).

6 Correlation and fringe fitting

6.1 General overview

The VLBI time delay $\tau$ is defined as the travel time of a plane radio wave which passes station 1 and then reaches station 2. At this point we are not going into detail how the reference point at each station is defined, but rather focus on the question how we can compare the recorded bit-streams from each station and extract quantities like group delay, phase delay, delay rate and signal-to-noise ration (SNR), respectively correlation amplitude. Before we have a look into theoretical aspects and study the actual correlation process, it might be good to reflect on the geometrical situation and the technical aspects of VLBI.
If we consider two VLBI stations distributed arbitrarily on the globe, we can easily identify the range of possible delays to be within

\[-\frac{R_e}{c} \leq \tau \leq \frac{R_e}{c},\]  

(28)

where \(R_e\) is the radius of the Earth and \(c\) denotes the speed of light. Thus, any VLBI delay has to lie between these boundaries which evaluate to approximately 21 ms. Without stating explicitly, we have assumed for this simple estimate, that the station clocks are synchronized to UTC sufficiently well. While this is not a straightforward task and requires the use of GNSS receivers, VLBI stations can normally synchronize their time-tagging units to UTC within ±1 µs. Revisiting the VLBI overview depicted in Fig. 1 one has to consider that the situation shown there is a static snapshot. During an actual VLBI observation, two main effects due to temporal variations need to be considered. First, the Earth is rotating during a scan is recorded. Thus, similar to GNSS, one needs to take into account a Doppler effect which impacts the data stream at each station in a way that the recordings suffer from delay rate effects. As similar but smaller contribution might also come from the frequency standard at each station, which very likely drifts from its nominal frequency and thus adds another delay rate-like effect. Ignoring higher order temporal effects, we can thus state that the correlator’s task to find the delay \(\tau\) between two stations can only be achieved if, at the same time, the delay rate \(\dot{\tau}\) between the recorded data streams is taken into account properly. This means, that the correlator’s task can be summarized as finding the most likely values for delay and delay rate that relate the sampling data from station 1 to those of station 2.

As will be addressed in more detail later, the correlation process is actually limited to a small search time interval applying a priori delays (to be re-added again at the end) from a priori geometry and clock offset information. Furthermore, a single observation is first decomposed in data chunks of individual periods of 0.1 to 1 s, so-called accumulation periods (APs), which later on are combined again. First, each sub-band channel is looked at separately and at the end these are combined as well leading to bandwidth synthesis (cf. Sec. 3).

### 6.2 Making use of the cross-correlation theorem

Ignoring for a moment that recorded signals suffer also from non-zero delay rate effects, and assuming that signals are captured with infinite temporal and spectral resolution we can write the cross-correlation function¹ as

¹ In radio astronomy \(\tau\) is often defined as the delay w.r.t. the second station, which would change the sign in the following equations
\[ C_{12}(\tau) = \text{Corr}(s_1, s_2) = \frac{1}{T} \int_{-\infty}^{\infty} s_1(t)s_2(t-\tau)dt. \]  

(29)

where \( s_1 \) and \( s_2 \) denote the continuous recordings at station 1 and station 2 Thompson et al (2007). Since we can only sample in discrete time steps and we should consider the geometrical constraints discussed before we have to collect a sufficiently large number of samples in order to be able to identify the correlation peak that corresponds to the time delay between the two recordings. We have to restrict ourselves also to a certain bandwidth \( B \) when we collect our samples which implies that we have to obey the Nyquist theorem which suggests that we should record one independent sample every \( \Delta t = 1/(2B) \). Thus, the finite integral from Eq. 29 turns into

\[ C_{12}[\tau] = \frac{1}{N} \sum_{k=1}^{N} s_1[k]s_2[k-\tau], \]

(30)

where \( \tau \) is considered to be a discrete (or integer) delay that has a granularity of \( \Delta t \). Before we discuss how we can achieve a sub-sampling delay resolution, we should look for efficient ways to compute the cross correlation function.

Considering that the integral in Eq. 29 looks almost like the convolution of two functions \( s_1 \) and \( s_2 \) which is defined as \( \int_{-\infty}^{\infty} s_1(t)s_2(\tau-t)dt \) and expressed in the form

\[ s_1 * s_2 = \mathcal{F}^{-1}(\mathcal{F}(s_1)\cdot\mathcal{F}(s_2)), \]

(31)

where \( \mathcal{F} \) and \( \mathcal{F}^{-1} \) are Fourier and inverse Fourier transforms.

### 6.3 Efficient ways to obtain the cross correlation and cross spectrum functions

In order to obtain the correlation function rather than the convolution between the two signals, we only need to flip the 2nd signal’s time order, i.e. changing from \( t-\tau \) to \( \tau-t \). However, this corresponds to a simple multiplication of the frequency axis with \( -1 \) in the Fourier domain. Furthermore, if \( S(f) \) is the Fourier transform of a signal \( s(t) \) then we are looking for \( S(-f) \). Fortunately, we can can make use of another very useful relation that states that \( S(-f) = \overline{S(f)} \) where long bar indicates the complex conjugate operator.

Thus, we have found a very simple and elegant way to express the correlation function \( C_{12}(\tau) \) by two Fourier transforms, a multiplication of Fourier coefficients and a single inverse Fourier transform, i.e.

\[ C_{12}(\tau) = \frac{1}{T}\mathcal{F}^{-1}\left(\mathcal{F}(s_1)\cdot\overline{\mathcal{F}(s_2)}\right), \]

(32)
Since we deal with discrete and finite sampling data we can replace the continuous Fourier transformation by Fast Fourier Transform (FFT) operations and obtain the cross correlation function by

\[ C_{12}[\tau] = \frac{1}{N} \text{IFFT} \left( \text{FFT}(s_1) \cdot \overline{\text{FFT}(s_2)} \right). \]  

(33)

Please observe that the cross correlation function is now implicitly defined as a complex function and its absolute value would correspond to Eq. 29 in case \( s_1 \) and \( s_2 \) were real-valued data streams. As we will see later, the original raw sampling data are of course real valued, but down-conversion or signal preprocessing steps can be preformed easier when dealing with complex-valued data streams. Often we are not after the cross correlation function, but try to access the so-called cross-spectrum between the two recordings,

\[ C_{12}[f] = \text{FFT}(s_1) \cdot \overline{\text{FFT}(s_2)}, \]  

(34)

which can be turned into the cross correlation function at any point by applying an inverse Fourier transform. We could have also obtained the cross spectrum from Eq. 30, if we apply the Fourier transform on the cross correlation function, i.e.,

\[ C_{12}[f] = \text{FFT} \left( \frac{1}{N} \sum_{k=1}^{N} s_1[k] s_2[k - \tau] \right). \]  

(35)

Thus, we have found two ways, summarized also in Fig. 19, to perform the correlation respectively the cross spectrum between the two signals \( s_1 \) and \( s_2 \). If we follow the concept of Eq. 35 we denote this choice as XF correlation, which indicates that the cross correlation is performed first in the time domain and then the Fourier transform is applied. If we, on the other hand, follow what is implicitly expressed in Eq. 34 we perform two FFTs first and then multiply in the spectral domain, which is the reason why this choice is called FX correlation. In the end the cross-spectra will be the same, but as it
turns out, FX correlation is preferable for nowadays VLBI operation due to the fact that FFT operations scale by $O(N \cdot \log_2(N))$ and thus outperform the classical XF approach already for a very small number of lags.

In order to understand why the cross spectrum appears to be the main target in correlation rather than the actual correlation function, we need to reflect on the fact that we are dealing with sampling data rates of several mega-samples per second (MSps) and thus cannot correlate the whole length of the data stream, but we will process that data in shorter pieces, so-called accumulation periods (APs), which in the ideal case even matches order-of-two FFT sizes and thus lead to a higher performance of FX type correlators. As long as we preserve coherency among those batches of spectra we can simply add them, i.e. integrate them and then perform the inverse Fourier transform on the stacked spectrum. Moreover, having access to the complex spectra in each of these batches allows us to compensate for delay and delay rate effects in a computational efficient way so that coherency can be preserved rather straightforward. Since the relation between the cross spectral characteristics and the corresponding correlation function is crucial for understanding the overall correlation process, Fig. 20 illustrates three simple cases that allow the reader to reflect on the concept of correlation. For the sake of simplicity we have created a short 64 sample data set of purely complex random data. From this data we created a second data set which was circularly rotated by 5 lags. Thereafter, both data streams were Fourier transformed and the cross spectrum was obtained following Eq. 34. The first row in Fig. 20 contains three plots depicting the corresponding cross spectrum characteristics (amplitude and phase) and the normalized cross correlation function. As for the latter, we can see that the correlation peak appears exactly at a lag distance of 5, which is what we would have expected based on how we have generated the second data set. We can also observe that the phase of the FFT points linearly varies with frequency. Thus, we have graphically confirmed the crucial relation between correlator phase $\phi$ and the delay $\tau$, which is usually expressed as

$$\phi = 2\pi \nu \tau. \quad (36)$$

This means that any delay between the two data stream manifests itself as a phase slope in the cross spectrum. In general, we can state that a phase slope in the cross spectrum will lead to a peak in the cross correlation function. Equipped with this knowledge we can now study two other situations, the impact of noise and the effect due to finite bandwidth. If we add a certain amount of noise to the second data set and perform the same analysis as before we obtain slightly different results as depicted in the middle row of Fig. 20. As expected, the phase slope in the cross spectrum becomes more noisy, which is not a surprise since we are no longer comparing only time shifted data but need to consider that the data-sets are now corrupted by the addition from a certain noise contribution. The other effect which we can observe
is that the correlation amplitude is no longer one as in the case of identical 
samples, but has decreased significantly. If we define the ratio between the 
correlation peak and the mean of the correlation function without that peak 
as the signal-to-noise ratio (SNR) we have gained a very intuitive under-
standing of how noise actually impact the SNR. Since a large SNR relates to 
a clear peak detection it is obvious that SNR will later (see Sec. 7) be used 
as a proxy to describe the uncertainty or precision of a delay measurement.

If we are not only considering noise contribution, but take into account 
that we are bounded to sample only over a very limited bandwidth, we can 
carry out another simulation (last row in Fig. 20) that depicts this effect very 
illustrative. Reducing the bandwidth can be easily simulated by setting the 
complex cross spectrum outside a certain pass band to zero and then perform 
the inverse Fourier transform to access the cross correlation function. When 
doing so, we can observe that the correlation peak is still around the 5 lags 
where we would have expected it to be, but we can also observe that the 
correlation peak is no longer identified by a sharp peak but is now located on 
the highest point of what can be approximated well be a quadratic function. 
In addition, the peak height has further decreased and together with our 
observation that the peak is now surrounded by a broader range of large 
values it is again very intuitive that we can deduce a relationship to SNR.
We can now directly relate to what we have defined in Eq. 6, where $T_{sys}$ is an equivalent measure of the noise contributions and $\Delta \nu$ represents the effective bandwidth of the receiving system. Lower noise (or lower system temperature) leads to a higher correlation peak and thus to a higher SNR. A wider bandwidth, together with a longer observation time, leads to a clearer phase slope in the cross spectrum and thus a higher and narrower peak in the cross correlation function which in the end translates to a higher SNR or a better formal error of the delay observable.

6.4 Fractional sample delay, fringe rotation and quantization noise

The correlator uses a-priori delay and delay models which allow to account for changes in the delay by dynamically adapting the read address of the raw data streams. The delay is tracked by the correlator using steps in sample units, which causes a difference between a priori delay and the discrete delay realized in the correlator. This difference must not exceed one sample unit within an integration/accumulation period, which normally ranges from one to a few seconds. The difference between the applied discrete delay and the actually modeled delay is referred to as fractional sample delay $\epsilon$ and is stored together with the cross spectrum for each integration period. Since we are dealing with cross spectra, we can easily correct for such delays by

$$C'_{12}(\nu) = C_{12}(\nu)e^{-i2\pi\nu(\epsilon_1-\epsilon_2)},$$

(37)

where $\epsilon_1$ and $\epsilon_2$ are the fractional delays at stations 1 and 2. Again, we see the advantage of using the cross spectrum since the operation expressed in Eq. 37 is computationally not very costly and can be implemented very easily. As shown in Fig. 21 one can implement the fractional delay correction directly after the FFT transformation of each station’s data stream.

When the delay rate has a positive value the telescope is steadily ”moving away” and thus the Doppler shift causes the signal received to be at lower frequency. Following that same logic, we can state that negative delay rates correspond to Doppler shifts towards higher frequencies. This implies that the signals from two stations cannot be correlated unless the frequency of the signals is referring to the same reference frequency. As depicted in Fig. 21, this can be achieved by frequency conversion of the raw sampling data and is easily implemented by multiplication of the data stream with a periodic signal that has a frequency which corresponds to the Doppler shift$^2$. This process is called fringe stopping, and is accomplished by multiplying cosine and sine functions with respect to the time-series data of the stations, i.e. before the

$^2$ This is a very simplified explanation. For more information the reader is referred to literature dealing with the actual implementation in more detail
data are either correlated or Fourier transformed. Since geodetic VLBI deals with observation frequencies of 8 GHz and above, the delay acceleration $\ddot{\tau}_g$, respectively the Doppler rate, needs to be taken into account as well when applying this correction. Fringe stopping is usually performed at frequencies calculated at the baseband frequency $\nu_0$ or at the center of the band $\nu_0 + \nu_B/2$ whereas the baseband is defined as the frequency of the radio frequency signal which is finally converted to the zero frequency. If the a priori model is sufficiently accurate, the resulting residual fringe rate will be at the level of a few mHz.

Beside the compensation for finite delay representation and fringe rotation one needs to consider that the recording systems are restricted by their data rate, measured in bits per second (bps). During analog to digital conversion the signal is represented by samples having $M$ quantization levels. Thus, it is obvious that the example discussed before and depicted in Fig. 20 would give different correlation amplitudes if we deal with quantized data. However, if the sampling rate and the quantization level is known, which is usually the case for VLBI we can correct the obtained correlation amplitudes for these discretization effect and obtain an unbiased correlation amplitude and thus an SNR value that refers to the original analog signal. According to Van Vleck and Middleton (1966) one can for example correct the correlation amplitude $\rho_c$ obtained from data sets which were 1-bit quantized by

$$\rho = \sin \left( \frac{\pi}{2} \rho_c \right). \quad (38)$$

Corrections for other sampling rates or quantization levels can be found e.g. in Thompson et al (2017).
6.5 Fringe fitting and bandwidth synthesis

Correlation is performed so that cross spectra are stored after a certain integration period, which leaves us with a time series of spectra. Knowing now that the inverse Fourier transform would give us the correlation peak at each of these integrations, we can now also try to evaluate how the position of this peak varies over time, which in the end will give us the delay rate value that corresponds to the duration of that particular scan. Thus, the process called fringe fitting, will use a set of consecutive cross spectra and try to find the most likely values of $\tau$ and $\dot{\tau}$ that maximize the correlation function, both in time as well as in frequency. Plotting all possible combinations of delay and delay rate gives us the so called 2D delay resolution function (see Fig. 22) which allows us to identify the most likely values for these two parameters.

For the delay direction we already know that we do not need to manually try all delay values but simply make use of the Fourier transform to obtain the one dimensional delay resolution function. As it turns out, we can also make use of the Fourier transform in the direction of consecutive cross spectra and obtain a peak where we have the most likely value for the delay rate. Thus, the 2D delay resolution function can easily be obtained by 2D FFTs performed on a stack of coherently integrated cross spectra. Once we identified the peak we can do a fine search around that location by a quadratic interpolation and obtain values of $\tau$ and $\dot{\tau}$ with a resolution that is better than the corresponding FFT resolution.

However, as we are not only collecting data from a single narrow channel, but record sampling data from several channels we are interested in using all this information together and obtain the most likely values of delay and
delay rate which maximize a correlation function that is derived from all these channels. This procedure of combining data from different frequency channels is referred to as bandwidth synthesis Rogers (1970).

Before we start, however, we have to correct for inter-channel phase differences stemming from different electric path lengths between the different in all channels. For this, we make use of the phase calibration signals which were injected at the receiving system (Sec. 5.4). In other words, we apply a phase shift to each of the channel’s cross spectra and then perform a fringe fit over all bands.

Finding the multi-band delay can be done in two different ways. The first one is that in a first step the cross-spectral function (amplitudes and phases w.r.t. frequency channel) is derived from individual channels as described for the single-band case. Then, a Fourier transform is carried out to convert the multi-band cross-spectral function into a multi-band delay function. The search for the delay basically works in the same way as for the individual channels, i.e., first a rough search for the location of the peak and then a suitable interpolation with selected spectral points in the vicinity of the peak Whitney (2000).

The second way of finding the final delay is a joint fit of the single- and multi-band delay as well as the delay rate in a common process as realized, e.g., in the PIMA software Petrov et al (2011). As we have seen before a wide effective bandwidth increases the SNR and improves the delay precision, which is the ultimate goal to achieve mm to cm precise VLBI delays. However, one needs to reflect on the fact that channels are usually not next to each other in the VLBI frequency bands but are separated by at least several MHz. Thus, instead of one large sinc-like correlation peak we will expect the actual correlation maximum to be surrounded by a certain number of slightly lower peaks, to which one refers to as side-lobes (Fig. 23). In general, the channel setup is carefully chosen so that one can expect minimum side lobe level in the actual data (cf. Sec. 3).

A further complication arises from the fact that the full bandwidth is not covered with one continuous channel and the set of channels do not align to each other without gaps in the frequency domain. For these reasons, we now face the challenge that our obtained delay can be biased by an integer multiple of the base ambiguity spacing. We can thus imagine that the graph Fig. 23 repeats every 100 ns which is the base ambiguity spacing for the frequency setup of this example.

To address the ambiguity issue in a bit more detail, we can review the tasks of single- and multi-band delay search also from the aspect of finding a phase slope that connects all the correlation phases of the individual channels. In the case of single band delays, we deal with such a fit in a straightforward way (see discussion earlier) as we have phase values at each of the FFT points and thus should be able to detect any phase slope rather easily if the SNR is high enough. However, in the case of multi-band delay search we have to connect the phases over a much wider bandwidth and the fact alone that
we have to deal with gaps between the individual bands, where we have no phase information at all, makes it already harder to find the actual phase slope that corresponds to the residual group delay. On top of that, there is a manifold of possibilities for possible delay slope values, due to the nature of phase measurements being limited to lie between $-\pi$ and $\pi$. Thus, we can only determine the residual multi-band delay by satisfying phase slope criteria, but need to consider that our obtained delay value might be off by several integer multiples of the base ambiguity spacing.

As described in more detail in Takahashi et al (2000) one can find a clear relation between the minimum channel spacing and the size of that base ambiguity. In general, one can state that the ambiguity spacing $\tau_{amb}$ corresponds to the greatest common divisor of all channel spacings $\Delta\nu_{max}$ by

$$\tau_{amb} = \frac{1}{\Delta\nu_{max}}. \quad (39)$$

Considering for example the eight USB channels of the standard geodetic X band setup (Sec. 5.3), we find a $\Delta\nu_{max} = 20$ MHz. Thus we face an ambiguity spacing of 50 ns, which is fortunately rather easily being dealt with given that the a priori information in the data analysis is good enough to provide theoretical delays with an uncertainty that is not exceeding half of the ambiguity spacing, i.e. 25 ns or about 7.5 meter.
7 Data analysis

7.1 Functional model

Data analysis of geodetic and astrometric VLBI analysis follows the same rules as any other geodetic problem with a surplus of observations for the determination of the parameters of interest. Let’s first look at the purely geometric relationships following the principal VLBI equation (Eq. 1) which has to be expanded to take into account the fact that the baseline vector \( \mathbf{b} \) and the unit vector in source direction \( \mathbf{k} \) have to be expressed in the same reference frame. The main cause for a more elaborate formulation is the Earth’s variable rotation Lambeck (1980). For this reason, the rotation matrices for precession and nutation, daily spin, and polar motion have to be applied.

Precession and nutation describe the motion of the Earth’s axis of figure in space with the angular arguments \( X \) and \( Y \), which are the coordinates of Celestial Intermediate Pole (CIP) in the celestial frame Dehant and Mathews (2015). This part of the variable rotation is modeled in the rotation matrix \( \mathbf{Q}(X(t), Y(t)) \). The Earth’s phase of rotation is measured as the so-called stellar angle \( \theta \) between the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO) Capitaine et al (2000). It should be noted here that originally the intermediate origins, which strictly speaking are rather zero-meridians, were called ephemeris origins. This was changed by a resolution of the International Astronomical Union (IAU) in 2006 (B2).

To make the stellar angle a time quantity, a conventional relationship was defined as

\[
\theta(T_u) = 2\pi \cdot \left(0.7790572732640 + 1.00273781191135448 \cdot (T_u - 2451545.0)\right) \tag{40}
\]

with \( T_u \) = Julian UT1 date (IAU Resolution 2000 B1.8). Solving for \( T_u \) provides UT1 from stellar angles \( \theta \) measured by VLBI observations (in fact the only means to measure UT1). \( \theta \) is the argument of the rotation matrix of the daily spin \( \mathbf{S}(\theta(t)) \). Finally, the variability of the Earth’s axis of figure w.r.t. the Earth’s crust is parameterized by the two components \( x_p \) and \( y_p \) of polar motion contained in the rotation matrix of this wobble effect \( \mathbf{W}(x_p(t), y_p(t)) \).

These three rotation matrices have to be applied to either transforming the baseline vector in the terrestrial reference frame into that in the celestial reference frame or the unit vector in the celestial reference into that in the terrestrial frame. Both transformations are equivalent and can also be applied in part to the radio sources and in part to the baseline vector, as is done in practice (see below). In a closed formula, they are embedded in

\[
\tau(t) = t_2 - t_1 = -\frac{1}{c} \mathbf{b} \cdot \mathbf{W}(t) \cdot \mathbf{S}(t) \cdot \mathbf{Q}(t) \cdot \mathbf{k} \tag{41}
\]
with

$$b = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix}$$  \hspace{1cm} (42)$$

for the Cartesian coordinates of telescopes 1 and 2, and

$$k = \begin{pmatrix} \cos \delta_c \cdot \cos \alpha_c \\ \cos \delta_c \cdot \sin \alpha_c \\ \sin \delta_c \end{pmatrix}$$  \hspace{1cm} (43)$$

for the position of the quasars in the celestial system (Sec. 1.2).

For the full expansion of the observed delay, we have to look at the relativistic contributions to the functional model and at several further effects to consider. Depending on the nature of these effects they are either applied as delay corrections $\Delta \tau$ or as corrections to the coordinate vector of the radio telescope $x_i$. It should be noted that at first we only consider those contributions, which can be applied a priori, while under Sec. 7.5.1 we address remaining model components, which can only be estimated.

### 7.2 Relativistic VLBI model delay

At this point relativistic effects need to be taken into account for proper modeling of the delays. The first item to consider is the transformation of time scales. We have to distinguish proper time measured at an arbitrary location and coordinate time which refers to the barycenter of a coordinate system. Any clock on Earth, first of all, measures proper time, while we will need time as coordinate time of the solar system barycenter. The proper time at a telescope is normally given in the UTC time scale (Universal Time Coordinated) which relates to the International Atomic Time TAI (Temps Atomique International) through an integer number of leap seconds $T_{LS}$ to keep UTC within 0.9 s of the astronomical time UT1. Terrestrial Time (TT) is an ideal time which a clock would indicate on the geoid. It has the same rate as TAI but an offset of 32.184 s originating from a changeover from Ephemeris Time (ET) to TAI, thus

$$TT = TAI + 32.184 \text{ s} = UTC + T_{LS} + 32.184 \text{ s}.$$  \hspace{1cm} (44)$$

The coordinate time of the Earth is Temps Coordonnées Géocentric (TCG) or Geocentric Coordinate Time. TCG [in fractions of Julian date] relates to TT [in fractions of Julian date] Bangert et al (2017) through

$$TCG = TT + L_G \cdot (T_{JD} - TT_0)$$  \hspace{1cm} (45)$$
with $T_{JD}$ being TT expressed as Julian date, $TT_0$ being 2443144.5003725 (i.e., TT at 1977 January 1.0 TAI), and $L_G = 6.969290134 \times 10^{-10}$ Petid and Luzum (2010).

The coordinate time of the solar system barycenter TCB (Temps Coordonnées Barycentrique) is related to TCG through

$$TCB - TCG = \frac{1}{c^2} \left( \int^{t}_0 \left[ \frac{v_\oplus^2}{2} + U_{ext}(X_\oplus) \right] dt + V_\oplus \cdot (X - X_\oplus) \right)$$

(46)

where the vectors $X_\oplus$ and $V_\oplus$ denote the barycentric position and velocity of the geocenter. The vector $X$ is the barycentric position of the observer and $U_{ext}$ is the Newtonian potential of all of the solar system bodies apart from the Earth, evaluated at the geocenter Bangert et al (2017). $t$ is TCB and $t_0$ is 1977 January 1.0 TAI.

Now we turn to the differences in arrival times at the two radio telescopes. The theoretical time difference $d_{TT}$ can be considered as being the same as that of $d_{UTC}$

$$d_{TT} = d_{TAI} = d_{UTC} = (t_2)_{TT} - (t_1)_{TT}$$

(47)

From the proper time frame to the Geocentric Coordinate Time (TCG) frame, the time difference needs to be scaled by

$$d_{TCG} = \frac{d_{TT}}{1 - L_G}$$

(48)

with $L_G = 6.969290134 \times 10^{-10}$ as above. It should be noted here that the VLBI coordinate frame is in fact (only) related to the TT time scale ("TT-compatible") for consistency with the other observing techniques in the ITRF as decided in an ITRF workshop in 2000, although IAU and IUGC resolutions call for TCG-compatibility Petid and Luzum (2010).

The next step is the transformation of the Earth-fixed geocentric coordinates of the radio telescopes, e.g., given in the ITRS (Sec. 1.2), $x_i$ into the barycentric frame $X_i$. For this, $x_i$ has to be transformed from ITRS into the Geocentric Celestial Reference System (GCRS) which has the same origin but fixed axis directions with respect to the extra-galactic radio sources of the ICRS. The transformation then just consists of adding the barycentric radius vector of the geocenter $X_\oplus(t_1)$ to the radius vector of telescope $i$ $x_i$ (in GCRS).

$$X_i(t_1) = X_\oplus(t_1) + x_i(t_1)$$

(49)

The reference epoch is $t_1$, i.e., when the signal arrived at telescope 1. All quantities to be addressed further down will also be referred to this time epoch because this epoch is the reference time of observation. Here it is in the TCG time scale which needs a scaling as in Eq. 45.

A very important contribution to the VLBI time delay originates from the retardation of the signal paths through gravitational bending of the celestial bodies. It is part of the relativistic formulation of the delay (Eq. 57). For
modeling the bending effect, the time of closest approach of a signal originating from a quasar and passing a celestial body is needed. For the \( J \)th gravitating body, it can be expressed as the minimum

\[
t_{1,J} = \min \left[ t_1, t_1 - \frac{K \cdot (X_J(t_1) - X_1(t_1))}{c} \right].
\]  

(50)

\( K \) is the unit vector from the solar system barycenter to the source in the absence of gravitational or aberrational bending Petit and Luzum (2010). This minimum actually means that it stays at \( t_1 \) if the signal does not pass the body before it arrives at telescope 1.

For the final expression, we need the vectors between the positions of the two telescopes and the gravitating body at the time of nearest approach, \( R_{1J}(t_1) \) and \( R_{2J}(t_1) \):

\[
R_{1J}(t_1) = X_1(t_1) - X_J(t_{1,J})
\]  

(51)

\[
R_{2J}(t_1) = X_2(t_1) - \frac{V_\oplus}{c} (K \cdot b) - X_J(t_{1,J})
\]  

(52)

\( V_\oplus \) is the barycentric velocity of the geocenter and \( b \) is the geocentric baseline vector.

The general relativistic delay, \( \Delta T_{\text{grav}} \), a.k.a. Shapiro term, then is

\[
\Delta T_{\text{grav}} = (1 + \gamma) \frac{G M_J}{c^3} \ln \left| \frac{R_{1J} + K \cdot R_{1J}}{R_{2J} + K \cdot R_{2J}} \right|
\]  

(53)

with \( M_J \) being the rest mass of the \( J \)th gravitating body and \( \gamma \) the so-called light deflection parameter equal to 1 according to the Theory of General Relativity (GRT).

The gravitational delay due to the Earth at the picosecond level is

\[
\Delta T_{\oplus} = (1 + \gamma) \frac{G M_\oplus}{c^3} \ln \left| \frac{x_1 + K \cdot x_1}{x_2 + K \cdot x_2} \right|
\]  

(54)

with \( M_\oplus \) being the rest mass of the Earth. The total gravitational delay can be expressed as the sum over all gravitating bodies, i.e., the Sun, Earth, the Earth’s moon and the other planets,

\[
\Delta T_{\text{grav}} = \sum_J \Delta T_{\text{grav}}^J
\]  

(55)

In the barycentric frame, the vacuum delay is

\[
T_2 - T_1 = -\frac{1}{c} K \cdot (X_2(T_2) - X_1(T_1)) + \Delta T_{\text{grav}}
\]  

(56)
Eq. 56 is now converted (back) by a Lorentz transformation from the barycentric into the geocentric system. This is done w.r.t. two groups of quantities, the barycentric vectors $X_i$ into the corresponding geocentric vectors $x_i$ and the related transformations between the barycentric time difference $(T_2 - T_1)$ and the geocentric time difference $(t_2 - t_1)$.

The geometric delay then follows

$$
(t_2-t_1)_{vac} = \frac{\Delta T_{grav}}{c} - \frac{K \cdot b}{c^2} \left[ 1 - \frac{(1 + \gamma)U}{c^2} - \frac{|V_{\oplus}|^2}{c^2} - \frac{V_{\oplus}w_2}{c^2} - \frac{V_{\oplus}b}{c^2} \left( 1 + \frac{K V_{\oplus}}{2c} \right) \right].
$$

(57)

This is the formulation for the geometry in vacuum Petit and Luzum (2010). $w_i$ is the geocentric velocity vector of telescope $i$, i.e., w.r.t. the geocenter. $V_{\oplus}$ contains the $x$-, $y$-, $z$-velocity components of the geocenter itself. $U$ is the total gravitational potential of the solar system at the geocenter neglecting the effects of the Earth’s mass. This formula thus contains the effects of daily and annual aberration. In the context of VLBI, this means that while the signal continues traveling for a time $\tau$ after arriving at telescope 1, telescope 2 has changed its position w.r.t. that at $t_1$ both in the geocentric and in the solar system barycentric frame.

### 7.3 Atmospheric refraction

Another very important contribution to the delay is that of atmospheric refraction. In terms of refraction, the atmosphere is separated into the charged compartment, generally called the ionosphere, and the neutral part which some authors call the troposphere. This, however, falls short of the fact that only about 75% of the refraction effects in the neutral atmosphere take place within the troposphere while 25% happen above Elgered (1993).

For observations in the radio frequency domain the ionosphere is dispersive. For this reason, VLBI observations have always been carried out with two separate frequencies (at X band and S band, Sec. 3) for calibration of the primary observing frequency, i.e., X band. The correction $\Delta \tau_{ion}^X$ for the X band observable is easily computed by

$$
\Delta \tau_{ion}^X = (\tau_S - \tau_X) \cdot \frac{\nu_S^2}{\nu_S^2 - \nu_X^2}
$$

(58)

where $\tau_S$ and $\tau_X$ are the measured multiband group delays and $\nu_S$ and $\nu_X$ the observed frequencies at S band and X band, respectively. Since the observed frequency of each band actually results from a combination of several frequencies through bandwidth synthesis, these frequencies need to be computed as a weighted mean over all channels $N$. 
\[ \nu_{gr} = \frac{\sqrt{\sum_{i=1}^{N} \rho_i \left( \sum_{i=1}^{N} \rho_i (\nu_i - \nu_0) \right)^2 - \left( \sum_{i=1}^{N} \rho_i (\nu_i - \nu_0) \right)^2}}{\sum_{i=1}^{N} \rho_i (\nu_i - \nu_0) \cdot \sum_{i=1}^{N} \rho_i - \sum_{i=1}^{N} \rho_i \cdot \sum_{i=1}^{N} \rho_i \frac{\nu_i - \nu_0}{\nu_i}} \]  

(59)

\( \rho_i \) are the weights of individual channels which can be computed from the number of correlated bits or SNR. \( \nu_0 \) is the reference frequency of the bandwidth synthesis setup, e.g., channel \#1 Herring (1983); Petrov (2001).

A common way of modeling refraction in the neutral atmosphere is the separation of \( \Delta \tau_{atm} \) into a hydrostatic and a water vapor part. The latter is often called the wet part. Both are separated further into a component reflecting the refraction effect in zenith direction and a mapping function for the conversion into the direction of the observation. The delay contribution \( \Delta \tau_{atm}(\varepsilon) \) is commonly represented as extra path length \( \Delta L_{atm}(\varepsilon) \) which is easily transformable to a time unit quantity through the speed of light. If the model only considers a dependency on elevation angle \( \varepsilon \), it can be composed as

\[ \Delta L_{atm}(\varepsilon) = \Delta L_{z}^h \cdot m_h(\varepsilon) + \Delta L_{z}^w \cdot m_w(\varepsilon). \]  

(60)

\( m_h(\varepsilon) \) and \( m_w(\varepsilon) \) are the mapping functions which are applied to map the extra path lengths or their time equivalents in zenith direction onto any non-zenith elevation angle \( \varepsilon \). A simple \( 1/\sin \varepsilon \) scaling is not sufficient, in particular at lower elevations because the curvature of the atmosphere as a consequence of the curvature of the Earth’s surface would not be taken into account. Most modern mapping functions thus are a function of \( \varepsilon \) and of a finite continuous fractional form with coefficients \((a, b, c)\) Herring (1992). They are specific for the hydrostatic (Eq. 61) and the wet component (Eq. 62) because the thickness of the effective atmosphere is more than 10 km in the first case and only about 2 km in the latter Niell (2000).

\[ m_h(\varepsilon) = \frac{1 + \frac{a_h}{b_h}}{\frac{1 + c_h}{\sin \varepsilon} + \frac{a_h}{b_h} + \frac{c_h}{\sin \varepsilon}} \]  

(61)

\[ m_w(\varepsilon) = \frac{1 + \frac{a_w}{b_w}}{\frac{1 + c_w}{\sin \varepsilon} + \frac{a_w}{b_w} + \frac{c_w}{\sin \varepsilon}} \]  

(62)
While in earlier mapping functions the coefficients mostly depended only on the geographic latitude of the telescope Herring (1992); Niell (1996), modern mapping functions are inferred from numerical weather models Böhm and Schuh (2004); Böhm et al (2006); Landskron and Böhm (2018).

The hydrostatic component of the extra path length in zenith direction $\Delta L_h^z$ follows a development of Saastamoinen (1972, 1973a,b), finally formulated by Davis et al (1985) as

$$\Delta L_h^z = 0.0022768 \text{[m hPa}^{-1}] \frac{P_o}{(1 - 0.00266 \cos 2\phi - 0.00028 H)}.$$ 

It predominantly depends on the barometric pressure at the height of the VLBI reference point $P_0$ [in hPa] and to a very small degree on latitude $\phi$ and the height above the geoid $H$. At sea level this corresponds to roughly $2.28 \text{ m}$ extra path length or $7.6 \text{ ns}$.

The hydrostatic path delay for each observation can be applied a priori according to

$$\Delta \tau_{atm}^h(\varepsilon) = \frac{1}{c} \cdot \Delta L_h^z \cdot m_h(\varepsilon). \quad (64)$$

A determination of the wet component along the line of sight $\Delta L_w$ with sufficient accuracy regularly fails because it would need a solution for an integral along the path length over the partial pressure of the water vapor $e$, the temperature $T$ and some compressibility factor $Z_w \approx 1$ Owens (1967) scaled by the mean temperature $T_m$ according to the model developed by Elgered (1982):

$$\Delta L_w = [1 + 6 \times 10^{-5} T_m] 3.754 \int_S \frac{e}{T^2} Z_w^{-1} ds. \quad (65)$$

The reason is that the water vapor in the atmosphere is highly variable in space and time.

Our inability to do a forward modeling of the wet component with sufficient accuracy leads to the generally applied strategy to just estimate an extra delay in zenith direction attributed to water vapor $\Delta \tau_w^z = \Delta L_w^z/c$ after the observations have been calibrated with a model hydrostatic contribution according to Eq. 64. More details on the background of refraction in the neutral atmosphere in VLBI can be found in Böhm et al (2013) while Alizadeh et al (2013) give detailed insights into ionospheric refraction.

### 7.4 Further model contributions

For the final geometric time delay in the terrestrial system, all known contributions have to be added to Eq. 57. The first ones are those of refraction through the ionosphere $\Delta \tau_{ion}$ according to Eq. 58 and through the neutral at-
mosphere $\Delta \tau_{atm}^b$ according to Eq. 64. At this point we also have to take into account that besides the aberration effects in vacuum embedded in Eq. 57, the path through the charged and neutral atmosphere at telescope 1 also permits telescope 2 to change its position due to the rotation of the Earth and its movement in space before the signal arrives there causing an additional aberration contribution $\Delta \tau_{abbs}^{atm}$

$$\Delta \tau_{abbs}^{atm} = \Delta \tau_{atm}^1 \cdot \frac{K \cdot (w_2 - w_1)}{c}.$$  

Further known corrections to be applied to the theoretical delay are tidal displacements and loading effects caused by the variability in the load forces of the oceans and the atmosphere. A complete list of them can be found in the Conventions of the International Earth Rotation and Reference Systems Service (IERS) Petit and Luzum (2010) with many references to the original scientific publications.

These effects all lead to displacements of the radio telescopes $\Delta X_{mod}$ on the inter-annual to sub-daily time scale. Since they are all in the linear domain for the analysis at hand, their impact can be easily transferred to the time delay domain through computing

$$\Delta \tau_{mod} = \frac{\partial \tau}{\partial X_{mod}} \cdot \Delta X_{mod}.$$  

It should be noted that the latter contributions, which are coordinate displacements in their original form, do not necessarily need to be applied at this stage. As adopted in some software packages they may also be applied to the a priori coordinates before going through the full relativistic formulation in Eq. 57 (Fig. 24).

The formulation of the final theoretical or a priori delay depends on whether the corrections mentioned above are actually applied to the observed delay, to the geometry or to the a priori delay. In

$$\tau_{apriori} = \tau_{vac} + \Delta \tau_{abbs}^{atm} + \Delta \tau_{E.T.} + \Delta \tau_{P.T.} + \Delta \tau_{O.L.} + \Delta \tau_{A.L.}$$  

the contributions $\Delta \tau_{E.T.}$ of the Earth tides, $\Delta \tau_{P.T.}$ of the pole tide, $\Delta \tau_{O.L.}$ of ocean loading, and $\Delta \tau_{A.L.}$ of atmospheric loading are applied to the theoretical delay. In contrast to this, $\Delta \tau_{T.E.}$ the contribution of the telescopes’ thermal expansion, $\Delta \tau_{G.D.}$ of the telescopes’ gravitational deformations, and $\Delta \tau_{Ins}$ of the telescopes’ cable delays (Sec. 5.4) are applied as corrections to the observed delay

$$\tau_{obs\_corr} = \tau_{obs} - \Delta \tau_{ion} - \Delta \tau_{atm}^b - \Delta \tau_{T.E.} - \Delta \tau_{G.D.} - \Delta \tau_{Cable}.$$  

All the $\Delta \tau$ have of course to be computed as differences related to telescope 1 and 2. Applying the respective effects to the theoretical delays provides the $c$ (computed $\equiv \tau_{apriori}$) vector and the vector of observations $o$ (observed
Forming the differences produces the linearized observation vector

$$y = o - c$$  \hspace{1cm} (70)

for the parameter estimation.

### 7.5 Parameter estimation

Various estimation schemes exist to determine the parameters of interest from VLBI multiband group delay observations. The primary difference is how these handle the stochastic nature of the clocks and in particular the atmosphere. Kalman Filter Kalman (1960); Herring et al (1990); Nilsson et al (2015); Soja et al (2015), least-squares collocation Titov (2000), square-root information filters Bolotin (2000) treat the wet part of the atmosphere and also the clock behavior as a purely stochastic process. However, most common are standard least-squares adjustments in Gauß-Markov models. It should be mentioned here, that in comparison campaigns it could not be substantiated that any of these estimation processes produces superior results. A drawback of the Filter solutions is, that combination of the results on the basis of normal equation systems is not possible directly, and for least-squares collocation, the extraction of a partly resolved normal equation system is quite complicated.

### 7.5.1 Gauß-Markov model

The estimation process in a Gauß-Markov model can be followed easily in Fig. 24, where the raw observations produced in the correlation and fringe fitting process are first corrected for all known effects mentioned before including the hydrostatic part of the atmosphere. On the right hand side, the radio source positions are transformed into position of date by the precession/nutation rotations composed in the rotation matrix $Q(t)$ in Eq. 83. The a priori telescope coordinates on the right hand side are mostly those of the ITRF (Sec. 1.2) which are transformed to the epoch of the observations by applying the corrections for the station velocity components in the ITRF and the rotations from the catalog frame to the frame of date through the Earth rotation parameters, polar motion and the Earth's phase of rotation. Further corrections to the coordinates are the effects of the Earth tides, the pole tide, oceanic and atmospheric loading effects as mentioned above.

Starting point for the explanation of the parameter estimation in the Gauß-Markov model is the so-called observation equation based on the initial formulation in Eq. 41. Since the two clocks are never synchronized to the accuracy
needed, the geometric model is expanded for the respective, time dependent clock offset. The clock behavior is time dependent and we can formulate the clock contributions $\Delta \tau_{\text{clo}}$ as second order polynomials

$$
\Delta \tau(t)_{\text{clo}} = -(T^0_0 + T^1_1 \cdot (t - t_0) + T^2_1 \cdot (t - t_0)^2) + (T^0_2 + T^1_2 \cdot (t - t_0) + T^2_2 \cdot (t - t_0)^2)
$$

(71)

to be added to Eq. 41. $T^*_j$ are the parameters of the clock polynomial at telescope 1 or 2.
The contribution of the refraction caused by water vapor, which we like to estimate, is expressed as the product of a zenith effect and a mapping function (Sec. 7.3). In a simple form, we can then write the observation equation as

$$\tau_{\text{obs}}(t) = -\frac{1}{c} \mathbf{b} \cdot \mathbf{W}(x_p, y_p) \cdot \mathbf{S}(\theta) \cdot \mathbf{Q}(X, Y) \cdot \mathbf{k}$$

(72)

$$- (T_0^0 + T_1^1 \cdot (t - t_0) + T_2^1 \cdot (t - t_0)^2) + (T_0^2 + T_1^2 \cdot (t - t_0) + T_2^2 \cdot (t - t_0)^2) + \mathcal{M}_w^1 \cdot \mathbf{A}^1 +$$

$$+ \mathcal{M}_w^2 \cdot \mathbf{A}^2 + + \tau_{\text{corr}} + \epsilon$$

with:

- \(\mathbf{k}\) unit vector in source direction (Eq. 43) its three components \(k_i\)
- \(\mathbf{b}\) baseline vector (Eq. 42)
- \(c\) velocity of light
- \(\mathbf{W}\) rotation matrix for polar motion
- \(\mathbf{S}\) rotation matrix for the daily spin of the Earth
- \(\mathbf{Q}\) transformation matrix for precession and nutation
- \(T_j^i\) parameters of the clock polynomial at telescope 1 or 2
- \(\mathcal{M}_w^i\) mapping function of the wet part of the atmosphere at station 1 or 2
- \(\mathbf{A}^i\) zenith wet delays at station 1 or 2
- \(\tau_{\text{corr}}\) sum of all corrections
- \(\epsilon\) measurement noise.

In a least squares adjustment of VLBI observations, the clock parameters always have to be estimated because of the lack of synchronized clocks at far distant locations. All other parameters of interest such as telescope coordinates, Earth orientation parameters, radio source positions, and zenith wet delays may be included in the vector of unknown parameters \(\mathbf{x}\) more or less freely depending on the observation constellation.

The classical least squares adjustment in a Gauss-Markov model reads

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x} + \mathbf{v}$$

(73)

with the covariance matrix of the observations \(\Sigma_{\mathbf{yy}}\). Here, \(\mathbf{y}\) describes the \(n \times 1\) vector of the abbreviated observations (Eq. 70), \(\mathbf{x}\) is the \(m \times 1\) vector of the unknown parameters and \(\mathbf{v}\) is the \(n \times 1\) vector of the residuals Koch (1999). \(\mathbf{A}\) is the \(n \times m\) Jacobian or linearized design matrix containing the partial derivatives of the observations w.r.t. the parameters to be estimated.
\[ A(i,j) = \frac{\partial \tau_{\text{obs},i}}{\partial x_j}. \] (74)

\[ \Sigma_{yy} \] is the covariance matrix of the observations according to Eq. 78 and the inverse of the covariance matrix is the weight matrix \( P = \Sigma_{yy}^{-1} \). The vector of the estimated parameters \( \tilde{x} \) is computed with

\[ \tilde{x} = (A^T \Sigma_{yy}^{-1} A)^{-1} A^T \Sigma_{yy}^{-1} y \] (75)

minimizing the sum squared of residuals

\[ \mathbf{v}^T P \mathbf{v} \ldots \text{min} \] (76)

where the residuals \( \mathbf{v} \) can be computed from

\[ \mathbf{v} = A \cdot \tilde{x} - y, \] (77)

Koch (1999).

7.5.2 Stochastic model

The covariance matrix of the observations \( \Sigma_{yy} \) is deduced from the a priori variance factor \( \sigma_0 \) and the cofactor matrix \( Q_{yy} \)

\[ \Sigma_{yy} = \sigma_0 Q_{yy}. \] (78)

The co-factor matrix of the observations \( Q_{yy} \) regularly only contains diagonal elements, i.e., the variances of the observations \( q_{ii} = \sigma_i^2 \) which are deduced from the signal-to-noise ratio of the correlation process (Sec. 6.3). It has been found very early that these are too optimistic and that the global test of the \( \chi^2 \) factor resulting from the quotient of the variance factor a posteriori \( \tilde{\sigma}_0^2 \) and the variance factor a priori \( \sigma_0^2 \)

\[ \chi^2 = \frac{\tilde{\sigma}_0^2}{\sigma_0^2} \quad \text{with} \quad \tilde{\sigma}_0^2 = \frac{\mathbf{v}^T P \mathbf{v}}{n - m} \] (79)

forces \( n \) observations and \( m \) unknowns is nowhere near the required unity. For this reason, the most common way of handling this problem in operational analyses, i.e., for standard solutions of the IVS, is by inflating the variances from the correlator \( \sigma_{corr}^2 \) with an additive constant \( \sigma_{add}^2 \)

\[ \sigma_{adj}^2 = \sigma_{corr}^2 + \sigma_{add}^2 \] (80)

to achieve that the input variances to the adjustment \( \sigma_{adj}^2 \) match those a posteriori and produces the \( \chi^2 \) in Eq. 79 to become close to 1.
The deficit here is twofold. The first is that the error contributions are computed to make the $\chi^2$ in Eq. 79 unity, thus they are purely empirical and lack any physical explanation. The second deficit is that correlations between observations are ignored entirely.

Various attempts have been made to cope with this problem, also for other space geodetic techniques, but solutions tended to be computationally expensive and thus not suitable for operational application. Lately, a very promising approach was published by using turbulence theory as the driver for the correlations. Due to its extremely good results and even more so that the algorithm has favorable runtime costs, it is well suited for operational applications Halsig et al (2016).

7.5.3 Extension of the functional model

Over the years, it had been realized that the functional model as formulated in Eq. 72 could be improved in several areas. The first one to mention is the pseudo-stochastic treatment of the clock behavior beyond a pure second order polynomial. Here, the parameter model is expanded for a number of additional parameters primarily for increased time resolution using continuous piece-wise linear functions (CPWLF) or linear splines De Boor (1978) of time segments of one hour or less. For this purpose the final observation equation (Eq. 83) also contains elements of these splines in the form of

$$x(t) = \frac{x(t_i) - x(t_{i-1})}{t_i - t_{i-1}}(t - t_{i-1}).$$

for time segments from $t_{i-1}$ to $t_i$. The same also applies to the temporal resolution of the zenith wet delay for which the observation equation (Eq. 83) is expanded in lines 6 and 7.

So far, the functional model assumes isotropy and just contains the wet delay in zenith direction as one or more parameters to be estimated. However, this does not account for the anisotropy of refraction which may be caused, e.g., by the increased thickness of the atmospheric layers towards the equator. For this reason, MacMillan (1995) introduced gradients in the model

$$\tau_{\text{grad}}(t) = +M_1F^1_w(\varepsilon^1) \cdot \cot \varepsilon^1 \cdot [G_n^2 \cos \alpha^1 + G_n^1 \sin \alpha^1] + M_2F^2_w(\varepsilon^2) \cdot \cot \varepsilon^2 \cdot [G_n^2 \cos \alpha^2 + G_n^2 \sin \alpha^2]$$

It should be emphasized here, that any remaining unmodeled hydrostatic contribution will be compensated for in the estimated wet component and in the gradients. The same applies to inaccurate pressure monitoring at the sites. For this reason the absolute values of the wet zenith delays may not be representative or even utterly wrong, e.g., if they are negative.
Taking all the augmentations of the functional model into account, the full observation equation in a Gauss-Markov model then reads

\[
\tau_{obs}(t) = \frac{1}{2} b(t) \cdot W(t) \cdot S(t) \cdot Q(t) \cdot k
\]

\[
- (T_0^1 + T_1^1 \cdot (t - t_0) + T_2^1 \cdot (t - t_0)^2)
\]

\[
+(T_0^2 + T_2^2 \cdot (t - t_0) + T_2^2 \cdot (t - t_0)^2)
\]

\[
- (T_1(t_0) + \frac{T_1(t_1) - T_1(t_0)}{t_1 - t_0} (t - t_0) + \ldots + \frac{T_1(t_i) - T_1(t_{i-1})}{t_i - t_{i-1}} (t - t_i))
\]

\[
+ (T_2(t_0) + \frac{T_2(t_1) - T_2(t_0)}{t_1 - t_0} (t - t_0) + \ldots + \frac{T_2(t_i) - T_2(t_{i-1})}{t_i - t_{i-1}} (t - t_i))
\]

\[
+ M F_w^1 \cdot [A T^1(t_0) + \frac{A T^1(t_1) - A T^1(t_0)}{t_1 - t_0} (t - t_0) + \ldots + \frac{A T^1(t_i) - A T^1(t_{i-1})}{t_i - t_{i-1}} (t - t_i)]
\]

\[
+ M F_w^2 \cdot [A T^2(t_0) + \frac{A T^2(t_1) - A T^2(t_0)}{t_1 - t_0} (t - t_0) + \ldots + \frac{A T^2(t_i) - A T^2(t_{i-1})}{t_i - t_{i-1}} (t - t_i)]
\]

\[
+ M F_w^1 (\varepsilon^1) \cdot \cot \varepsilon^1 \cdot [G_n^2 \cos \alpha^1 + G_c^1 \sin \alpha^1]
\]

\[
+ M F_w^2 (\varepsilon^2) \cdot \cot \varepsilon^2 \cdot [G_n^2 \cos \alpha^2 + G_c^2 \sin \alpha^2]
\]

\[
+ \tau_{corr} + \epsilon
\]

(84)

with

\[ T_i^* \] parameters of the clock polynomial at station 1 or 2

\[ T_i^1(t_i) \] coefficients of linear splines at station 1 or 2 for a specific epoch \( t_i \) parametrized as continuous piece-wise linear functions (CPWLW), i.e., linear splines

\[ M F_w^* \] mapping function of the wet part of the atmosphere at station 1 or 2

\[ AT^*(t_i) \] zenith wet delays parametrized as CPWLW at station A or B for a specific epoch \( t_i \)

\[ G^*_n/c \] atmospheric gradients in north or east direction at telescope 1 or 2

### 7.5.4 Solution types

Before turning to the final estimation, we have to distinguish between single session analysis and multi-session, so-called global, analysis. The standard parameter set for single session solutions consists of four groups of parameters which are considered as local parameters, i.e., only valid at the epoch of the session:

- the parameters of the clock model
- radio telescope coordinates \( x_i, y_i, z_i \)
- Earth orientation parameters, i.e., the two polar motion components \( x_p, y_p \), the Earth’s phase of rotation UT1-UTC, two adjustment parameters
- for the precession/nutation model dX, dY
the parameters of the refraction model due to the wet component of the atmosphere.

The single session analysis always is the first step which is mainly devoted to preparing the session data for establishing the final setup of this session. First of all the ambiguities have to be resolved, which result from possible ambiguous phase slope (Sec. 6.5). Special care has to be taken that in all possible triangles the sum of the delays has to close to zero within a few nanoseconds.

After this has been done, the X band group delays are corrected for ionospheric refraction effects according to Eq. 58. It can be expected that systematic biases between the S and X band delays prevail which contaminate the ionosphere correction. However, since these are of a constant nature, they will finally disappear in the clock parameter estimates.

Next comes the parameterization of the clock and atmosphere parameters mainly depending on the number of individual delay observations and the possible and necessary time resolution of the parameters. Eq. 83 contains all the parameters which are estimated in routine processing. However, we have to consider two peculiarities. The first is that we can only estimate clock parameters which are relative to a reference clock. Regularly one of the most stable clocks is selected from experience. For this purpose the design matrix does not contain coefficients for the reference clock resolving the singularity which would appear otherwise.

The second caveat is the estimation of the atmospheric refraction parameters. Normally, VLBI telescopes are far enough apart that the estimates of the wet zenith path delays of any two telescopes are de-correlated. However, with the advent of more than one radio telescope at a single observatory, e.g., at Wettzell (Germany), Onsala (Sweden), or Ny Alesund (Norway), the partial derivatives of the first order wet zenith delay of two nearby telescopes

$$\frac{\partial \tau}{\partial \Delta T^1(t_0)} = M F^*_w(\varepsilon)$$

and of course all other continuous piece-wise linear polygons are highly correlated through the almost identical elevation angles. For this reason, only relative zenith wet delays can be estimated in cases where the two telescopes are close together. As soon as a third telescope provides observations on longer baselines with different elevation angles, the estimation of absolute zenith wet delays for all telescopes becomes possible again.

The estimation of parameters of piecewise linear functions does only work if a sufficient number of observations is available in each segment. Unfortunately, sometimes this is not the case because of failures or deliberate exclusions. In such a case the system of normal equations would become degenerate. However, this is regularly cured by introducing constraints in the form of pseudo observations to the effect that the rates between every two functional values should be zero.
\[ \frac{d(F(t_i) - F(t_{i-1}))}{dt} = 0. \] (86)

The selection of the variances of these observations determines how loosely or how tightly the constraint affects the estimation. For the clock parameters normally a standard deviation on the order of about $5 \times 10^{-14}$ s/s and for the zenith wet delay of about 50 ps/h is being chosen.

Another issue concerns the fact that the VLBI technique is only capable of producing relative telescope coordinates. This means that any solution setup per se is free of datum and the normal equation system would be singular if all telescope coordinates and the Earth rotation parameters ($x_p, y_p, UT1-UTC$) are tried to be estimated. It has a rank defect of six, three translations and three rotations. In a single session, the simplest option would be to fix the three coordinate components of one telescope and the three Earth rotation parameters, i.e., eliminate them from the normal equation system.

Another more flexible way being less dependent on the coordinates of a single telescope is the introduction of no-net-rotation (NNR) and no-net-translation (NNT) conditions for at least three non-collinear telescopes or even all of them in the network Angermann et al (2004). The selection of the datum telescopes depends on quality of the a priori coordinates being suitable for the datum definition at hand.

As a final step of the session-wise analysis, an iterative identification process of significant outliers in the observations can be added applying different kinds of outlier identification schemes Koch (1999). Thus outliers can then be down-weighted for exclusion from the parameter estimation.

The second type of solutions is a multi-session or global solution. These solutions are carried out to determine parameters which are valid/constant for the whole observing period. They serve to estimate

- radio source positions
- telescope coordinates at a reference epoch
- linear or higher order velocities of the radio telescopes due to geodynamic effects

Other experimental solutions can be conceived, e.g., for the estimation of the relativistic factor $\gamma$ or Love numbers.

A simple way of running these solutions is by stacking pre-reduced normal equation systems from the session-wise analysis. Pre-reduction in this context means that parameters are excluded from the solution vector by reducing the system of normal equations by Gaussian elimination steps Angermann et al (2004). When all local parameters are pre-reduced from the session’s normal equation system, it contains only the coefficients of the global parameters set up initially. New lines and columns, have to be entered for the linear or otherwise parameterized, e.g., harmonic, telescope motions Angermann et al (2004).
The radio source positions are mostly treated either as constant global parameters or as session-wise, so-called arc, parameters. The latter approach is chosen when a radio source evidently manifests position variations. However, new approaches also allow for intermediate modes with linear spline approximations at predetermined or automatically chosen time intervals Karbon et al (2017).

As the final step, the global parameters, as determined by the stacked normal equation system, are reinserted in the original session normal equation systems to also estimate the local parameters. With all parameters estimated the observation residuals can be computed and an assessment of the errors can be carried out.

The general concept of treating the datum in global solutions is the similar to that of single session solutions with NNR/NNT conditions Angermann et al (2004). In most of the cases, the latest realization of the ITRF Altamimi et al (2016) is taken as the datum frame. For the radio source positions, mostly the defining sources of the ICRF2 Fey et al (2015) are chosen as the celestial datum for new and/or improved radio source positions.

8 Results

Session-wise results are of particular interest in the case of Earth orientation parameters. Unique for VLBI are the results of the Earth’s phase of rotation UT1, which cannot be determined with such a quality and time resolution with any other technique in particular not with any satellite technique. The reason is that UT1 is directly correlated with the ascending nodes of any satellites’ orbits. The phase of rotation is reported as time difference of UT1 – UTC. Formal errors are at the level of 2-3 µs for contemporary network sessions while daily single baseline sessions of one hour duration only (so-called Intensives) provide UT1–UTC with a standard deviation of about 8-15 µs.

Fig. 25 Results of UT1-UTC between two epochs where leap seconds were introduced.
Unique are also the estimated adjustments to the precession/nutation model. They are accurate to about 80 $\mu$as and represent not so much inaccuracies in the precession/nutation model but variations caused by free core nutation Dehant and Mathews (2015) which are at the level of about 1 $mas$.

Daily polar motion estimates provide a reliable time series of the long term evolution of polar motion (Fig. 26). Although the formal errors are slightly worse than those of the IGS polar motion time series, they are needed for cross-validation.

From the time series of telescope coordinates, baseline lengths can be inferred. These are independent of any datum and can conveniently be used for stability investigations. One of the most prominent baseline is that between the 18 m telescope of MIT Haystack Observatory near Westford (MA) at the East coast of the U.S. and the 20 m telescope of the Bundesamt für
Kartographie und Geodäsie and the Technische Universität München near Wettzell, Germany (Fig. 27). It should be noted that the input to the time series is not constrained in any way and still the approximation by a linear regression is extraordinary precise. Fig. 27 has been produced by a Web tool of the IVS under http://www.ccivs.bkg.bund.de/EN/Quarterly/ VLBI-Baseline/vlbi-baseline_node.html where plots of all VLBI time series of baseline lengths are available interactively.

Results originating from VLBI global solutions guarantee consistency between parameters estimated in the same solution setup. Standard parameters here are radio source positions, which are being used in the construction of the ICRF Fey et al (2015). Telescope coordinates and velocities regularly enter the computations of the specific realizations of the International Terrestrial Reference System (ITRS), the ITRF Altamimi et al (2016); Bachmann et al (2016). The standard deviations of the coordinates of the most reliable and often used telescopes are at the level of about 1 – 3 mm with velocity components having standard deviations of 0.1 mm/y.

Acknowledgements Almost all of the section on correlation were written by Thomas Hobiger of Chalmers University of Technology, Gothenburg, Sweden. I am grateful for the help of a real specialist in this important part of VLBI. I also thank Klaus Börger and Armin Corbin for proofreading of the manuscript.

References


Dehant V, Mathews PM (2015) Precession, Nutation and Wobble of the Earth


stability of a radio telescopes reference point using a terrestrial laser scanner:
Case study at the onsala space observatory 20-m radio telescope. ISPRS Journal
of Photogrammetry and Remote Sensing in preparation
Jacobs CS, Arias F, Boboltz D, Boehm J, Bolotin S, Bourda G, Charlot P, de Witt
A, Fey A, Gaume R, Gordon D, Heinkelmann R, Lambert S, Ma C, Malkin Z,
Nothnagel A, Seitz M, Skurikhina E, Souchay J, Titov O, ICRF-3 working Group
(2014) The ICRF3 Roadmap to the next generation International Celestial Refer-
ence Frame. In: American Astronomical Society Meeting Abstracts #223, Ameri-
can Astronomical Society Meeting Abstracts, vol 223, p 251.25
of the foundation of the 20 meter radio telescope. Technical Report 178, Onsala
Space Observatory, Sweden
of Fluids Engineering 82(1):35–45, DOI 10.1115/1.3662552
extension of the parametrization of the radio source coordinates in geodetic VLBI
10.1007/s00190-016-0954-1
Lambeck K (1980) The Earth’s Variable Rotation
Landskron D, Böhm J (2018) VMF3/GPT3: refined discrete and empirical trop-
s00190-017-066-2
of radio telescope reference points with sub-mm accuracy: results from a cam-
paign at the Onsala Space Observatory. J GEODESY 87(8):791-804, DOI
10.1007/s00190-013-0647-y
Lössler M, Haas R, Eschelbach C (2016) Terrestrial monitoring of a radio tele-
scope reference point using comprehensive uncertainty budgeting in- vestiga-
tions during cont14 at the onsala space observatory. JoG 90(5):467-486, DOI
10.1007/s00190-016-0887-8
MacMillan DS (1995) Atmospheric gradients from very long baseline interfero-
tory observations. Geophysical Research Letters 22(9):1041–1044, DOI 10.1029/
95GL00887, URL http://dx.doi.org/10.1029/95GL00887
Matveenko LI, Kardashev NS, Sholomitskii GB (1965) Large base-line radio in-
http://dx.doi.org/10.1007/BF01038318
Moran JM, Crowther PP, Burke BF, Barrett AH, Rogers AEE, Ball JA, Carter
JC, Bare CC (1967) Spectral line interferometry with independent time stan-
dards at stations separated by 845 Kilometers. Science 157(3789):676–677, DOI
10.1126/science.157.3789.676, URL http://dx.doi.org/10.1126/science.
157.3789.676
Ungar Publishing Co., New York
Niell AE (1996) Global mapping functions for the atmosphere delay at radio wave-
95JB03048
Niell AE (2000) Improved atmospheric mapping functions for VLBI and GPS. Earth,
doi.org/10.1186/bf03352267


tinental baseline and the rotation of the earth measured by radio interferometry. Science 186:920–922
Index

accumulation period, 28, 31
ambiguity, 36, 37
astrometry, 5
azimuth, 12
bandpass, 11
bandwidth synthesis, 11, 25, 36
cable calibration, 27
Cassegrain system, 15
correlation, 27
cross correlation theorem, 28
cross spectrum, 30
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
delay calibration, 27
delay search, 35
Fast Fourier Transform, 30
feed horn, 15, 22
FFT, 30
Fourier transform, 29
fractional sample delay, 33
fringe rotation, 33
FX, 31
Gregorian system, 15
H-Maser, 11
ICRF, 8
International Celestial Reference Frame, 8
International Terrestrial Reference Frame, 6
International VLBI Service for Geodesy and Astrometry, 6
ITRF, 6
IVS, 6
mapping function, 13
Nyquist, 25
phase calibration, 26, 36
prime focus, 15
quantization, 33
quasar, 5, 7, 12
quasars, 8
radio source, 5, 7
SEFD, 10
Shapiro term, 41
SNR, 27
system temperature, 10
upper sideband, 25
USB, 25
XF, 31