

## An Ever more penetrating View

---

**Jan Noordam**

ASTRON

Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

E-mail: noordam@astron.nl

**Johan Hamaker**

ASTRON

Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

E-mail: hamaker@astron.nl

*50 Years Westerbork Radio Observatory, A Continuing Journey to Discoveries and Innovations*  
Richard Strom, Arnold van Ardenne, Steve Torchinsky (eds)

Jan Noordam\*

## *Between the astronomer and his result, stands software...*

An instrument like the WSRT makes measurements of astronomical objects. These include images, frequency spectra, polarization and time variability. Apart from nominal parameters like sensitivity and resolution, the quality of such “Views on the Universe” depends critically on our ability to calibrate the various instrumental effects. In our case, this is done primarily by means of software.

Over the last five decades, the quality of radio astronomical observations has been improved by several orders of magnitude, allowing astronomers to penetrate ever deeper into (the secrets of) the Universe. Throughout, the WSRT group has played a leading role in this.

## Introduction and Overview

In order to separate astronomical “truth” from instrumental effects, an accurate model of the instrument is required, and of the intervening propagation medium. Such a Measurement Equation (ME) contains many parameters, whose values have to be estimated with great accuracy. Especially if very faint objects are to be studied in the presence of very bright ones. Fortunately, the contaminating bright sources can be used to solve for ME parameters, and thus to calibrate the instrument.

One rather useful definition of calibration is “the capability to subtract the effects of bright sources from the data with (very) high accuracy”, so that faint objects may be studied. This capability implies a (very) accurate knowledge of the values of all instrumental parameters.

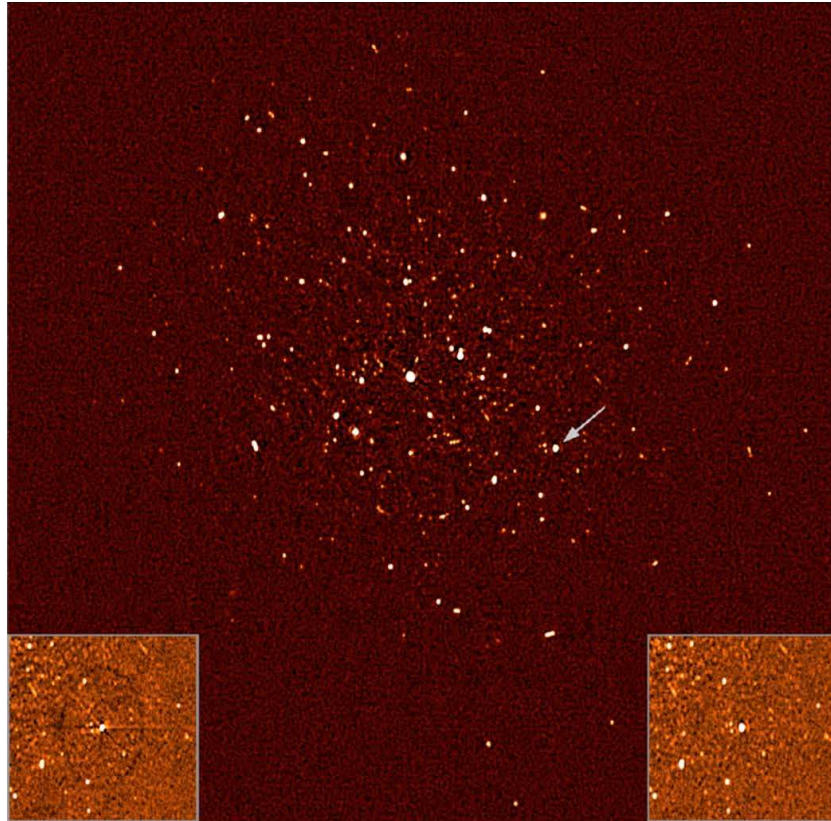
The development of calibration algorithms in radio aperture synthesis over the last 50 years has been remarkably successful. Apart from making many astronomical discoveries, it has put us in a position to contemplate the creation of much larger and more sensitive radio telescopes. On the other hand, we have not been particularly good at offering these algorithms effectively to the users of our telescopes, thus causing them to be under-utilised. This is a matter of some urgency.

---

\* ASTRON, The Netherlands

The WSRT group has played a leading role in the development of calibration algorithms. For instance, for most of the 50 years under discussion, the WSRT has held the world record of “Dynamic Range”. This widely used quality measure is the ratio between the flux of the brightest source in the field, and the faintest source that can be discerned. This ratio has improved by almost five(!) orders of magnitude, from about 100 in 1970 to 8.000.000 (8 million!) in 2017.

*3C147 (Smirnov/Perley) The Dynamic Range of radio images has improved by almost five(!) orders of magnitude, from about 100 in 1970 to 8.000.000 (8million!) in 2017. The WSRT group has played a major role throughout*



An important reason for the accuracy of WSRT calibration is that the instrument has been very carefully engineered, which leads to better starting conditions for the calibration process. In addition, the WSRT has a number of features that lead to a simpler Measurement Equation, with fewer parameters to be calibrated. These features include equatorial mounts, on-axis receivers, redundant spacings and correlators with essentially zero closure errors.

The superior calibration of the WSRT has produced much admiration, but only a few astronomers were prepared to invest the considerable extra effort required to get the most out of the instrument. But the exercise has put ASTRON in a position to take the lead in thinking about the next generation of giant radio telescopes, like LOFAR and SKA, for which such precision is no longer a luxury, but a necessity.

In addition to the development of calibration algorithms, the WSRT group has also played a leading role in efforts to make it easier for astronomers to process data taken with the WSRT and other radio aperture synthesis telescopes. This ranged from offering a WSRT Data Reduction Service during the 1970's, to trying to unify the proliferation of software packages that sprang up after the Selfcal revolution around 1980. These worldwide efforts have only had limited success, despite the obvious advantages.

### *The glorious first ten years (1970-1980)*

Radio astronomy is a relatively young subject, but tremendously important. 80% of the Nobel prizes awarded to astronomical subjects are in the radio regime. And the WSRT was precisely on time. For 10 years, it was the most sensitive telescope in this new and exciting wavelength window, where many important discoveries were made. In addition, there was not yet much competition from rival wavelength regimes, which required telescopes on mountain tops or in space.

The success of the WSRT was made possible by the fact that Prof Oort in Leiden had been one of the first to realize the potential of radio astronomy. As a result, considerable technical experience had been built up with the 25 m Dwingeloo telescope, led by a very thorough Lex Muller. But the cherry on the cake was a remarkable student called Wim Brouw, who not only extended the theory of aperture synthesis to help design the WSRT, but who also happened to understand the new but essential phenomenon of computers. Thanks to him, WSRT calibration and imaging fully operated from day one.



*Wim Brouw was one of the first to understand computers. As a student, he made sure that WSRT imaging and calibration fully operated at the opening in 1970. He later re-implemented the RSC software as NEWSTAR*

As a result, many talented people from all over the world flocked to the Netherlands to use this marvellous instrument, and to pick the abundant low-hanging fruit. One might say that the first decade was the most glorious period of the WSRT, even though the image quality was still relatively poor because calibration techniques like Selfcal had not yet been discovered. Fortunately, the CLEAN deconvolution technique had been invented by Jan Högbom in 1972. Arguably the high point of the WSRT was the 1977 Symposium in Groningen on “Image Formation from Coherence Functions in Astronomy”, which was attended by everyone who was anyone in radio astronomy processing.

After 1980, the WSRT remained at the scientific forefront for decades, despite being somewhat overshadowed by newer telescopes elsewhere, with more active PR departments than the modest Dutch were comfortable with. The WSRT also remained a technological leader, with cutting-edge correlators, frontend receivers, IF systems, observation modes (including VLBI, the Sun, and pulsars), time/frequency dissemination, shielding against interference, and calibration software. All these aspects are treated in detail elsewhere in this book. As a last hurrah, the WSRT is now pioneering the art of using receiver arrays in the focal plane of the dishes, which significantly widen the field of view.

### *Imaging by means of Aperture Synthesis*

WSRT is an aperture synthesis telescope. It consists of an array of 14 dish antennas, with a diameter of 25 m each, placed on an East-West line of 3 km length near the village of Westerbork, in the Netherlands.

An aperture synthesis telescope is an imaging instrument, just like an optical telescope. Both receive electro-magnetic radiation from objects in the sky through a given aperture (literally: opening). The resulting distribution of electric fields in the aperture gives rise to something called a “visibility function”, which can be turned into an image of the sky by means of a Fourier Transform (FT). In an optical telescope, the FT is performed by a lens or mirror that is placed in the aperture. Physically, it concentrates the light from different “stars” on different locations in the focal plane, thus forming an image of the sky.

The same principle can be used to make images of the radio sky, since optical and radio waves are both electro-magnetic waves, albeit with different wavelengths. The visibility function of the incoming radio waves may be sampled by means of an array of radio antennas, and then turned into an image by a performing the Fourier Transform with a computer. This was first demonstrated in a practical instrument by Ryle, who was awarded the Nobel Prize in 1974.

So radio aperture synthesis measures the Fourier Transform of the image, rather than the image itself. For this reason, it has been called “indirect imaging”, as opposed to “direct imaging” with optical telescopes. As mentioned elsewhere, indirect imaging has considerable advantages for calibration, and for building large telescopes.

The first radio telescopes consisted of single dishes, which were used for a time-consuming form of direct imaging, observing only one pixel at a time. Such dishes soon reached their mechanical limits, with a diameter of about 100m. Fortunately, indirect imaging with arrays of many modestly-sized dishes greatly increase the sensitivity and the spatial resolution.

The field of view of an aperture synthesis telescope is determined by the primary beam, i.e. the response pattern of a single element of the array. At an observing frequency of 1400 MHz (21 cm), the 25 m dish of the WSRT is sensitive to an area of half a degree on the sky, the size of the full Moon. The spatial resolution is determined by the overall size of the array, which can be as large as the 12000 km diameter of the Earth, or even larger.

### *Sampling the Visibility Function*

The visibility function (i.e. the Fourier Transform) of the sky may be sampled by correlating the signals from “interferometers”, i.e. pairs of antennas. Each interferometer represents a point in the “aperture plane”, or uv-plane. Its coordinates (u,v) are determined by the length of the baseline between the two antennas, and its orientation w.r.t. the observed sky (i.e. as it is “seen” from the observed source).

Because of Earth rotation, the projected length and orientation of the baseline changes slowly in time, so each interferometer samples the visibility function along a curved line (an ellipse) in the aperture plane. In this way, the equivalent of an aperture might be said to be “synthesized”. As soon as enough “visibilities” have been collected, they may be transformed into an image of the sky by means of the inverse Fourier Transform. Obviously, the image quality will be better if the aperture is more fully sampled, simply because more complete information is available. In that case, calibration is also easier.

The process of sampling the aperture plane can be greatly speeded up by using arrays of many antennas. The N=14 dishes of the WSRT contain  $N*(N-1)/2=91$  independent interferometers, i.e. combinations of two dishes. Obviously the antenna locations should be chosen in such a way that each interferometer samples a different ellipse in the aperture plane.

Most (10 out of 14) WSRT dishes are located at regular intervals of 144 m on a straight line. This was done because the WSRT was designed in the 1960's when computers were not yet very powerful, and certain imaging algorithms had not yet been discovered. Nowadays, such a configuration is not considered efficient because it does not maximize the number of independent interferometers. For instance, there are 10 baselines with the same length (144 m) and orientation, so they sample the same uv-point. On the other hand, these “redundant” spacings can be used for calibration (see below).



# Re-formulating Theory: The Measurement Equation and Polarimetry

Johan Hamaker

The front page from the famous paper by Hamaker, Bregman and Sault which introduced the Measurement Equation to aperture synthesis imaging.

The first experiments in radio-astronomical interferometry used pairs of identical antennas each with a receptor for a single component of the incoming radiation. Such interferometers can be represented by a simple algebraic equation relating the single output to the intensity of the incident radiation. An obvious way to double the sensitivity of the instrument is to add second receptors for the complementary component that the first receptor pair misses. The two signal pairs travel through separate channels through the interferometer and can be independently calibrated.

instrumental errors. The very large effect of ionospheric *Faraday rotation* must be treated separately. In all this, it is also tacitly assumed that all antennas and receptors are identical, invariable and mechanically tracking the motion of the source during an observation. This method has been coded in software packages and prescriptions for their use. They have been discussed at length in papers with page-wide formulas which only a very few specialists fully understand. For most others, polarimetry remained a black art whose fruits they happily consumed. Whether the rules were always correctly applied we cannot know, but in general the outcomes reported have been credible enough

Kurt Weiler added a twist that could be applied only in the WSRT designed with rotatable receptors. By rotating those of the movable telescopes at 45 degrees relative to the fixed ones (the '*crossed dipole*' configuration) he made calibration simpler and more robust. In this mode the WSRT observed successfully for a decade, with the world suspiciously watching. In the 1980s when, for the sake of higher sensitivity, the fixed-fixed and movable-movable interferometers were activated, the crossed-dipole mode had to be abandoned.

Many may have felt uneasy with the complexity of polarimetry and its inelegant treatment, yet the astronomical community accepted it because it produced plausible results – and anyway there was no alternative in sight. Meanwhile, maintenance of

this software was becoming an intolerable burden and the hard decision to make a fresh start from scratch was made at several places.

In 1991 the international *AIPS++* consortium was set up to develop a new universal software infrastructure that would serve the needs of all of astronomy once and for all. At ASTRON, inconclusive discussions were bypassed by Wim Brouw who single-handedly managed to develop (in old-fashioned Fortran) a complete system (later to be named NEW-STAR).

Both efforts stumbled on fundamental problems associated with polarimetry. Tim Cornwell found that *AIPS++* lacked a fundamental cornerstone which he called the *Measurement Equation*. Wim was bothered by ambiguities about the order in which multiple errors in interferometric polarimetry should be corrected. Johan Hamaker was horrified by the ugliness of customary polarimetric processing and decided that there must be a more proper way.

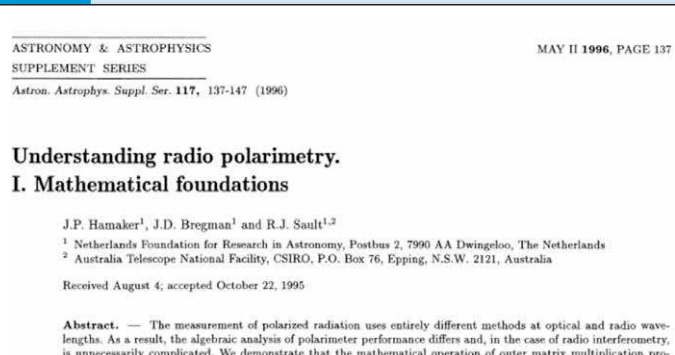
For half a century, radio astronomy had been built upon the scalar interferometer equation that ignores the two-component *vector* character of electromagnetic radiation. Instead of ordinary algebra, mathematics offers a tool to handle such vectors known as Linear Algebra, the algebra of *vectors* and *matrices*. In a breakthrough paper *Hamaker, Bregman and Sault* (1996) reformulated interferometer theory from scratch, showing how matrix algebra can represent all familiar features and quandaries of

the traditional theory in a simple and natural way.

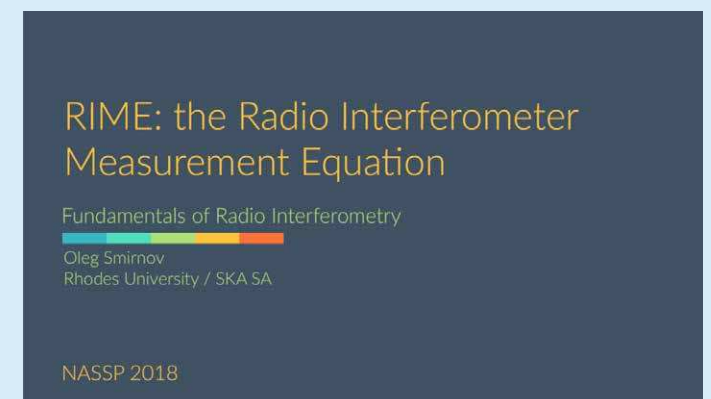
When Cornwell received a preprint, he immediately recognised the proposed matrix equation as the one that he had been groping for in vain and promptly orchestrated a change-over in *AIPS++*. Observing astronomers on the other hand rejected the 'matrix nonsense', claiming that *AIPS++* had better things to do.

Indeed the conversion job, though inevitable, came on top of a project that had been overloaded with ambitions and expectations from the beginning. A few years later it collapsed under the burden, but it has not lived in vain. Its legacy is a large body of well-designed software infrastructure (now known as CASA) which is properly maintained and upon which others can build further for their own purposes. Major beneficiaries are ALMA and our own LOFAR.

LOFAR with its mechanically static hardware cannot in any way be made to satisfy the demand that the antenna hardware tracks an observation. It is therefore fair to say that LOFAR as we know it could not have existed without the Measurement Equation. Conversely, the needs of LOFAR have enforced a software development that might otherwise have happened much more slowly. Replacing the very root of interferometer theory, the Measurement Equation represents a true revolution in radio astronomy. Since its introduction, an entire new generation of astronomers has been brought up who consider it as self-evident. Yet in practice they, too, must in many cases resort to legacy programs from the previous century. It is very fortunate that two major events came at the right moment to speed up modernisation: The establishment of *AIPS++* and the construction of LOFAR as prototype of a new generation of *phased-aperture* telescopes.



Many radio sources exhibit the astronomically important property of *polarisation* which produces a correlation between the two signal components. Similar correlations are spuriously produced by errors in the interferometer itself. To handle these effects, the pair of interferometer equations must be dissected and patched up with auxiliary equations that in turn must be made tractable through approximations. For most sources the polarisation effects are small enough to justify the approximations and the same is true for the

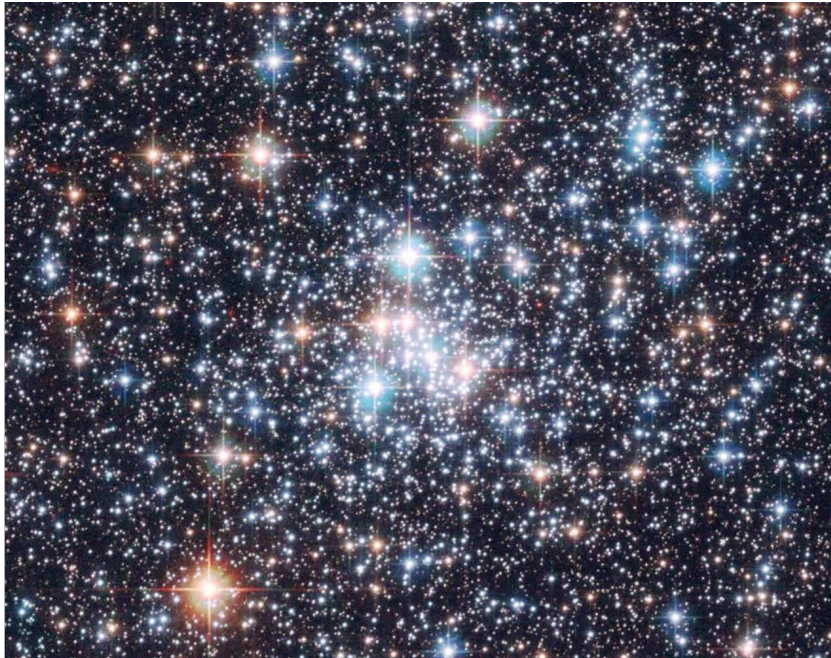


In real use: Oleg Smirnov teaching about the Measurement Equation in South Africa (courtesy: Oleg Smirnov, formerly ASTRON, now at SKA S.A.) See also his MOOC lecture: [https://www.youtube.com/watch?v=ipsyz9Ngv\\_g&index=9&list=PLVc4L-hESH-SokNtCwD4NWuQkkyzkFDSQL&t=0s](https://www.youtube.com/watch?v=ipsyz9Ngv_g&index=9&list=PLVc4L-hESH-SokNtCwD4NWuQkkyzkFDSQL&t=0s)

## Deconvolution

In all imaging instruments, a point source will not be imaged as a point. It will be imaged as a “blob”, which will be larger (i.e. the image will be fuzzier) if the aperture is smaller. In addition, an incomplete sampling of the aperture will cause such a blob to be surrounded (i.e. convolved) by a “point spread function” (PSF) which can be quite extended. The PSF around bright sources can contaminate the entire field-of-view.

An optical image taken with the Hubble Space Telescope illustrates how point sources are convolved with a PSF that is caused by the imaging instrument. In this case, it is cross-shaped, caused by the struts of the secondary mirror



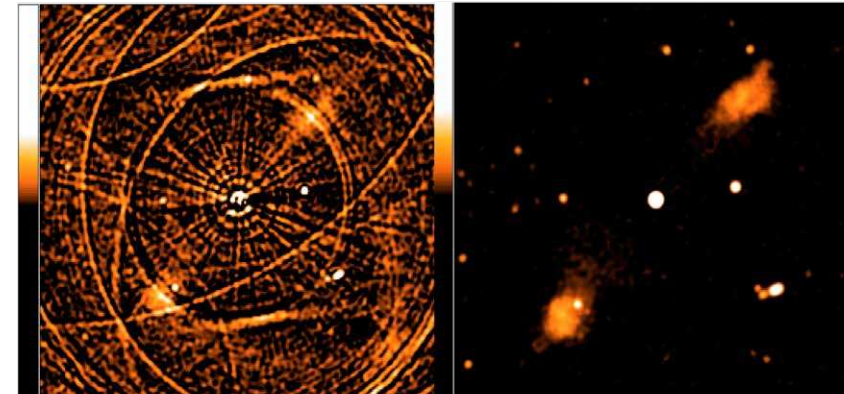
If the PSF of an imaging instrument is accurately known, it can be subtracted from the image, centered on the position of bright sources. This process is called “deconvolution”. The CLEAN algorithm developed by Jan Högbom in 1972 greatly contributed to the early success of the WSRT. Eventually it became clear that it is much more accurate to subtract bright sources from the visibility data, rather than from the pixelated image.

At the 1977 colloquium in Groningen, there was a fierce debate between Steve Gull and others about the most fundamental merit function to be used in a deconvolution technique called Maximum Entropy. This was settled by Jan Högbom, who showed that any convex function would work equally well.

During the design of SKA in the 1990's, there was a debate whether the collecting area should be distributed over a relatively small number of large array elements, or a large number of smaller ones. This was settled by the realization that fewer elements cause higher PSF sidelobes, giving rise to considerable PSF sidelobe noise caused by the many faint radio sources that cannot be detected for individual subtraction. Their number will greatly increase with the highly sensitive new telescopes. This problem can only be remedied by using arrays of many elements, and thus low PSF sidelobes.

## Calibration of aperture synthesis radio telescopes

The nominal PSF (determined by the aperture sampling pattern) will be distorted by instrumental errors, caused by the receiver electronics or the atmosphere. Therefore, the calibration of an aperture synthesis instrument may be defined as “the capability to subtract the PSF of (very) bright sources with (very) high accuracy”, clearing the view for faint objects in the field. For this, it is essential to estimate the instrumental errors very accurately.



A WSRT image of 3C236 before and after subtracting the Point Spread Function (PSF) of the synthesised beam from the “dirty” image (left) to produce the sharp image (right) in this case of the Seyfert galaxy 3C236. Note the rings and radial spikes caused by the WSRT PSF.

In terms of calibration, indirect imaging has a number of huge advantages. First of all, in direct imaging, all instrumental effects are mixed up in every pixel. But since a visibility sample only depends on the signal from two antennas, the main instrumental effects are conveniently separated in the data, which makes it much easier to determine their values (see also Selfcal below). Secondly, bright sources may be subtracted much more accurately from visibility samples than from a pixelated image. Thirdly, the electronic measurement noise on visibility samples affects the PSF in a different way than the photon noise in direct images.

## Three categories of instrumental effects

As explained above, aperture synthesis measurements are made by interferometers, i.e. pairs of antennas (e.g. dishes). So, each visibility sample depends on only two antennas, and contains information about all the sources in the field-

of-view. In this context, we can distinguish three categories of instrumental errors/effects:

1. IBE: Interferometer-based effects are associated with individual interferometers. They can be either additive (thermal noise) or multiplicative (correlator errors).
2. ABE: Antenna-based effects are associated with individual antennas. Examples are cable-lengths, electronic gain, atmospheric phase, etc.
3. DDE: Direction-Dependent effects depend on viewing direction, i.e. they vary over the field-of-view. Examples are (differences in) antenna beam-shapes (including pointing), ionospheric phase, etc.

These categories are not mutually exclusive, i.e. an instrumental effect can belong in more than one category. Most effects also dependent on time and frequency.

The key to calibration is the reduction of the number of independent parameters that have to be solved for. In any case, the number should be considerably smaller than the number of visibility samples. The huge success of Selfcal (see 2GC below) is based on the assumption that all instrumental effects are ABE, so in an array of  $N$  antennas there are only  $N$  independent errors for every set of  $N*(N-1)/2$  interferometers. Since the advent of digital correlators, which do not generate their own IBE, we may assume that this “Selfcal condition” is largely satisfied. Except of course for thermal noise, so Selfcal does not work for low  $S/N$ .

The number of independent parameters can also be reduced considerably if it can be assumed that they are roughly constant, or vary only smoothly, in time, frequency or direction. In that case, one may solve for the coefficients of smooth functions.

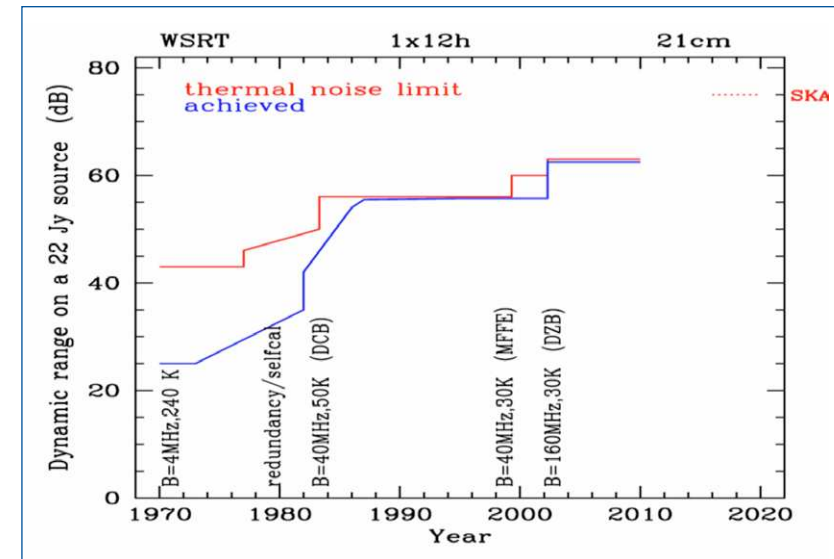
Note that visibilities can only be corrected for one direction in the sky, so “corrected visibilities” do not exist in the presence of DDE. Therefore, DDE corrections can only be applied during gridding, which greatly increases the processing requirements.

#### Four Generations of Calibration

In the development of radio astronomical data reduction, we may (roughly) distinguish four Generations of Calibration (GC):

- 1GC: Reliance on instrument stability (1960)
- 2GC: The Selfcal Revolution (1980)
- 3GC: Direction-Dependent Effects (DDE, 1995)
- 4GC: Statistical analysis of the residuals (2010)

After the Selfcal revolution around 1980, the calibration kept up with the increasing sensitivity of the WSRT (see also Chapter 10).



After the Selfcal revolution around 1980, the calibration kept up with the increasing sensitivity of the WSRT. Thanks to the “unnecessary” attention to detail of the WSRT group, we might be in a position to calibrate the much more sensitive new telescopes.

In the following sections, the four generations of calibration will be discussed in some more detail.

#### 1GC: Reliance on instrument stability

The measured Visibility samples are complex numbers, each with an amplitude and a phase. To first order, the instrumental errors caused by the receiver electronics or the atmosphere can also be expressed as a complex number, with which the “true” Visibility is multiplied. During the first decade, the WSRT was calibrated by observing a bright point source with a known flux and position. For each interferometer, any deviations from the expected phase (zero) and amplitude (the calibrator flux) were interpreted as the instrumental error for that pair of antennas.

These instrumental errors were assumed to be constant over the 12 hours of the subsequent observation of the desired source. For the most common observing frequency of 1400 MHz (a wavelength of 21 cm), this assumption was accurate to a few degrees of phase, and a few percent of amplitude. The result was a dynamic range of about 100, i.e. the residuals between the nominal and actual (distorted) PSF for the brightest source in the field were about 1% of the peak flux of that source.

As mentioned above, 1GC was good enough to pick the low-hanging fruit in the glorious first decade of the WSRT. Meanwhile, there were serious concerns about the phase stability of the new Very Large Array (VLA) in the USA, which had longer baselines and operated at higher frequencies. Plans were made for calibrating much more frequently than every 12 hours. Fortunately, that proved not to be necessary.



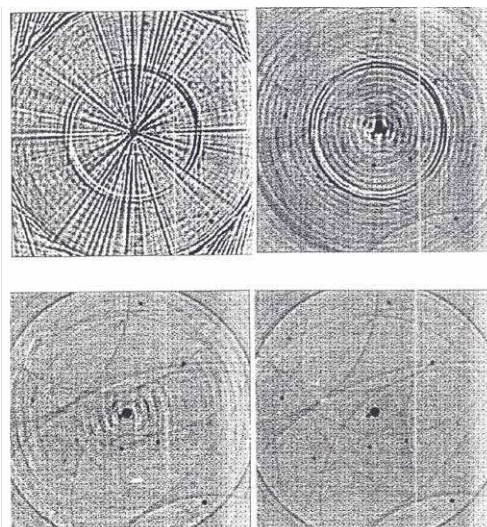
## 2GC: The Selfcal Revolution (1980)

The analog electronics that were used to correlate the signals from the two antennas of an interferometer caused a significant “closure error” of its own. This changed with the introduction of digital correlators in the late 1970s. From then on it could be assumed that the instrumental error of an interferometer was simply the product of the (complex) errors of its two antennas (the Selfcal Condition). This meant that the  $N*(N-1)/2=91$  samples measured simultaneously by an array of  $N=14$  antennas shared only 14 independent errors. This reduction of the number of free parameters made it possible to “self-calibrate” the instrument continuously during the observation.

*Four stages of the calibration of a SRT field that contains the bright source 3C48, made with the wonderful imaging software of 1980. Since the nominal PSF of the WSRT is corrupted by instrumental errors, it can only be properly subtracted if these errors are known accurately. If not, faint sources in the field cannot be studied because the entire image will be contaminated with the remains of the PSF of 3C48*

In a well-designed aperture synthesis telescope, the only closure error, and thus violation of the Selfcal Condition, is the thermal noise, which is different for each interferometer. This will only affect the calibration solution if there are no bright sources in the field, i.e. the S/N ratio is low. And even then the problem is minimised because random noise will not cause any systematic effects.

Selfcal starts with an initial model of the observed field, e.g. a single point source. For each timeslot (e.g. 1 minute) the values predicted by the model are compared to the actually measured visibilities of all 91 interferometers. The differences are then used to solve for the instrumental errors of the 14 antennas. After correcting the measured data for the estimated antenna errors, an improved model of the observed field is derived, perhaps including some of the fainter sources. This iterative process is repeated until the residuals between predicted and measured data are small enough.



Of course this process may appear a bit like Baron Munchhausen pulling himself out of the swamp by his own hair. It converges because the sky is “empty” in the sense that the observed field contains only a limited number of bright sources, which are well separated from each other. It also helps that we may assume that the sources have positive flux, and that they must be within the primary beam. Obviously, Selfcal will converge more quickly (and reliably) if the number of antennas is larger. The minimum seems to be about  $N=6$ .

Selfcal was invented by the VLBI community because the phases on the very long baselines involved were too unstable for 1GC. It was initially

missed by the WSRT group because the measured data were (for excellent reasons) organised by interferometer rather than by timeslot. The WSRT version of Selfcal was developed when Noordam started using the redundant baselines of the WSRT array for calibration (see below), which required organization by timeslot.

Selfcal caused a worldwide revolution. It made VLBI possible, it saved the VLA, and everywhere the image quality improved by leaps and bounds. The best results were obtained by the WSRT, partly because the telescope was more carefully (over-)engineered than its competitors, and partly because the WSRT group was a bit more obsessive about it. During the 1980s, more than 20 software packages were written to exploit this wonderful new technique. About 4 have survived to this day: AIPS, MIRIAD, DIFMAP, NEWSTAR (see also below).

NB: Since interferometers essentially measure phase differences between widely separated points in the incoming wave front, the absolute position of a source is lost. Therefore, observations are always relative to bright calibrator sources with a known position (and flux).

## Redundant Spacing Calibration (RSC)

RSC was the WSRT road to the discovery of Selfcal. If two interferometers are redundant, i.e. if they have the same baseline length and orientation, they should yield the same visibility value, irrespective of the observed brightness distribution. Any differences must be caused by instrumental effects. This can be exploited for calibration if it may be assumed that all instrumental effects are associated with the  $N$  antennas in the array, rather than the  $N*(N-1)/2$  different interferometers, i.e. that the Selfcal Condition is satisfied. As mentioned above, this became the case with the Digital Line Backend (correlator) built by Albert Bos in the late 1970's.

RSC by itself produced perfect “scans” of  $N*(N-1)/2$  complex visibility samples that had been observed simultaneously. Since comparison of redundant visibilities is insensitive to absolute phase and amplitude, different scans still had to be “aligned” to each other to produce a perfect image. This was done with a model of the observed brightness distribution, just like Selfcal.

It was soon realized that RSC should be seen as a (model-independent) extra constraint on the Selfcal solution. The joint solution is easily implemented, but it is not much used in practice.

It should be noted that RSC only works well if all antennas “see” the same sky. This is not the case when the antennas have different beamshapes. Therefore, RSC will only be of limited use for calibration of LOFAR or SKA



### 3GC: Direction-dependent effects (DDE)

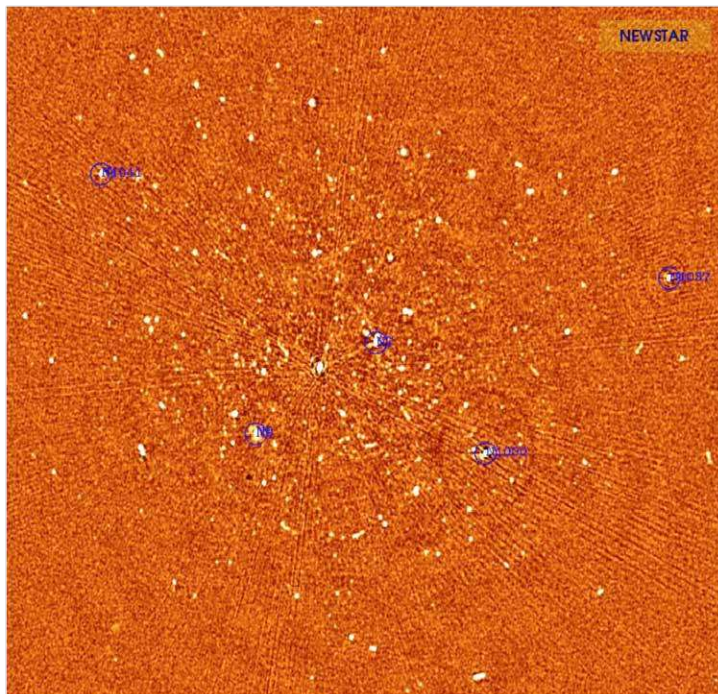
During the first decade or so, Selfcal was used to solve for a single complex error per antenna, which was assumed to be valid for the entire field of view. This worked well enough for fields that were dominated by a single bright source. Since the estimated errors were most valid for the position of that source, it could be subtracted very effectively, making all the fainter sources much more visible.

But as telescopes became more sensitive, it became a problem that the response beams of the N antennas were slightly different, so that antenna errors varied differently as a function of position (direction) in the field. The ionosphere also causes significant “direction-dependent effects” (DDE), especially at observation frequencies below 400 MHz.

During the 1990s, selfcal was generalized to take account of DDE. For instance by solving for different antenna errors separately for other moderately bright sources in the field. Or by assuming that errors vary smoothly over the field, and solving for the coefficients of low-order smooth functions.

Since uv-data can only be corrected for a single direction, the existence of DDEs meant that corrected uv-data do not exist. This meant that it became much more difficult and expensive to generate calibrated images of the sky. The only way to apply DDE is as part of gridding the uv-data onto a regular grid, prior to Fourier Transforming them to make an image.

*3rd Generation Calibration (3GC) deals with Direction Dependent Effects (DDE), i.e. instrumental errors that not only differ per antenna, but also vary smoothly over the field-of-view. In this WSRT image, the bright source 3C147 has been completely removed from the centre by 2GC, but there are still remnants of the PSF of other sources that are affected by DDEs. This problem gets worse with increasing telescope sensitivity and field-of-view.*



The generalization of Selfcal was extended also to the (smooth!) variations of antenna errors as function of time and frequency. Smoothness can usually be assumed for time variations, but not always for variations over the frequency band. Estimating smooth functions requires solving for fewer parameters, so it is less expensive in processing. And of course, it is only possible to solve for (much) fewer parameters than the number of available data.

NB: If the field is dominated by a single bright source, good results may be obtained by assuming that the response beams of all dishes are identical (2GC). But for the best results it is necessary to take into account the small differences between the beams (3GC).

### 4GC: Statistical analysis of the residuals

After 3GC has successfully removed the PSF of all sources brighter than 3-10 times the instrumental noise, the image residuals still contain many treasures, and problems. They must be approached by statistical means, which is the realm of a new kind of astronomer, who is much more mathematical. They must try to extract the signature (if not the image) of the EoR from the deviation of the noise from a Gaussian distribution.

The Final Frontier is the PSF sidelobe noise of the sea of ever fainter sources that are left in the image after 3GC. The number of those sources increase with faintness, and at some point they start to flow together because of confusion. It would be nice if there were some way to remove the sea of faint sources from the data without knowing their individual positions and fluxes, but with just their overall statistics. Unfortunately, that probably violates the Second Law of thermodynamics.

## Observing in more Dimensions

Much of the discussion above is about imaging the radio sky, which is after all the main observing mode of an aperture synthesis radio telescope. Over the years, the WSRT developed into a most superior imaging instrument, through careful engineering, and excellent software. After the Selfcal revolution in 1980, the image quality kept up with the increasing sensitivity brought by better receivers and (much) wider observing bandwidth.

### Spectral studies

Images are all very nice, but most (80%) of the astronomical information is in the frequency spectrum. It contains information about the velocity of astronomical objects, and of their temperature and chemical composition. Therefore, the default mode of operation is to making a 3D image cube in which the observed field is imaged for a range of different frequencies. Calibrating the frequency response of the telescope is a serious complication, especially since this response may vary over the field of view.

In the case of the WSRT, standing waves between the dish and the focus box caused the baseline of the frequency response to be sinusoidal, with a typical period of 17 MHz. The latter was determined by the length of the support legs of the focus box, and the observing wavelength. This effect became more serious as the total observing bandwidth increased over time, from 2 MHz to 100 MHz or more. Fortunately, the standing waves have disappeared with the introduction of focal plane arrays which do not reflect radio waves like the old single-pixel feeds.

### *Time variability*

On the whole, the Universe is pretty constant on human time-scales, and most of the time variations are caused by instrumental effects. When astronomical objects do vary in time, they are invariably point-like. This makes it easy to distinguish their variations from instrumental effects. This is greatly helped by the fact that the variations are only in amplitude, and that the WSRT has excellent amplitude stability.

The most rapid, and intensively studied, time-variations are those of pulsars. Since the 1990's, the WSRT has dedicated hardware (PuMa) to search for regular subsecond variations in the raw signals from the various antennas, and in which signals are processed differently e.g. because of de-dispersion required to add the pulse structure and for increased sensitivity.

### *Polarization*

The WSRT has excellent polarization properties. Most of the instrumental polarization is eliminated by having circular dishes and on-axis receivers. There are also advantages to having equatorial mounts because the dishes do not rotate w.r.t. the sky during observations.

Polarisation calibration is difficult because there are not many bright calibrator sources with linear polarization in the sky, and even fewer with circular polarization.

## **Serving the users of our radio telescopes**

### *The WSRT Data Reduction Service*

Radio aperture synthesis telescopes like the WSRT were the first to use computers for observation and data processing and it was deemed too difficult for astronomical users in the early 1970s. Moreover, few astronomers had experience with indirect imaging, i.e. observing the Fourier Transform of the radio sky, rather than the image itself. And finally, the WRST was made available to anyone who put in a worthy observing proposal, rather than just to the expert owners. For all these reasons, it was decided that WSRT data reduction would

be done by experts, using programs written largely by Wim Brouw. The users would receive science-ready data products, like images.

Although this was hailed as an enlightened policy, it was soon met by cries of ingratitude. First of all, since the WSRT was a unique and revolutionary telescope with great discovery potential, the impatient users fretted about the delay of several weeks before they received their images. Even then, the Service personnel was not perceived to be particularly responsive to polite inquiries, muted complaints or constructive suggestions. To be fair, these officials were not really in a position to take effective action, for instance by reducing the data in a different way, or by modifying the software themselves. The situation did not improve when the Service was moved from Leiden (where most of the early users were) to Dwingeloo.

In the early 1980s, another model was pioneered by a new competitor of the WSRT, the recently finished Very Large Array (VLA) in the USA. The Americans gave the users their raw visibility data and a software package (AIPS) and told them to do the rest themselves. This worked much better, since the AIPS developers were moderately responsive to feedback, and the users could also write their own software in AIPS.

Eventually, the WSRT developed and distributed its own data reduction package (NEWSTAR, see below), and many WSRT users also used other packages like AIPS or MIRIAD to process WSRT data. At the same time the WSRT User Service continued operation until well into the 1990s.

Of course the whole issue of Service Data Reduction will surface again in the present day. The reason is that the data volumes that are produced by the new generation of radio telescopes are thought to be too large to be taken home by the user, who in any case will not have a suitable computer (or software) to process it. The problem may be solved by developments in computer networks and software. In any case: "Those that do not learn the lessons of History, are doomed to repeat it".

### *The NEWSTAR data reduction package*

During the 1970s, WSRT observations were transported to Leiden by means of 40 MB tapes, to be processed on the mighty IBM360 computer. Amazingly, the staff in Westerbork or Dwingeloo were not supposed to make images or even inspect the data. Fortunately, Johan Hamaker developed calibration and imaging software to test some ideas of his own. This was used gratefully by Jan Noordam to investigate the use of the redundant spacings of the WSRT for calibration. The rapidly improving image quality then attracted the attention (and hard work!) of Ger de Bruyn, a highly respected and knowledgeable WSRT user, to whom this book is dedicated. The result of this collaboration was the WSRT version of Selfcal, which was described in a Nature paper (Noordam, de Bruyn, 1981) after the Dynamic Range exceeded 1:10.000, on 3C84 @6cm.



# A picture is worth a million image points: Per 5 aspera ad astra

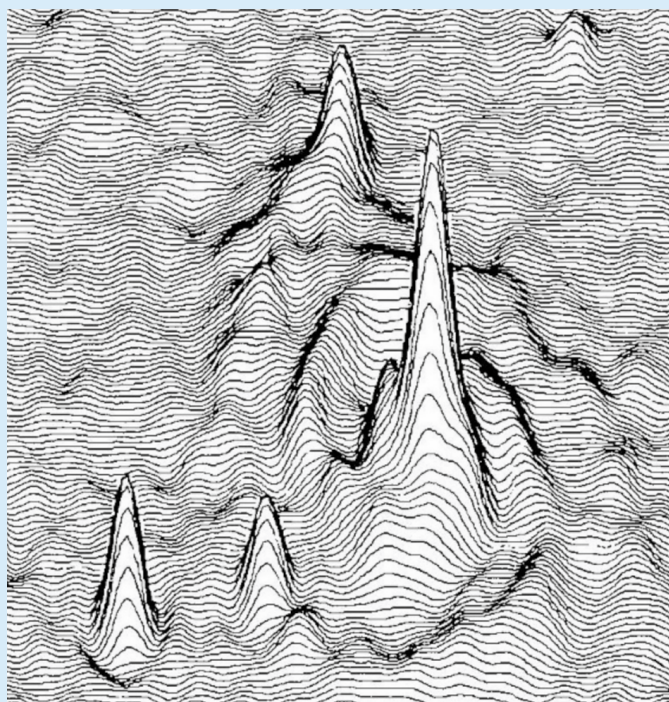
Johan Hamaker

From its very beginning astronomy has been based on visual information. Even today our visual abilities are unsurpassed for recognising scientific patterns in the flood of new observations. Ever since this flood took the form of *rasters* of numbers in computers the quest for suitable visual representations has been on. How can we translate such an array into an image that we can *look* at? With modern technology this may seem a trivial question to ask, but in the 'dark ages' before the utopia of fast high-quality monitors and printers became a reality for every household, we had to resort to much more primitive techniques.

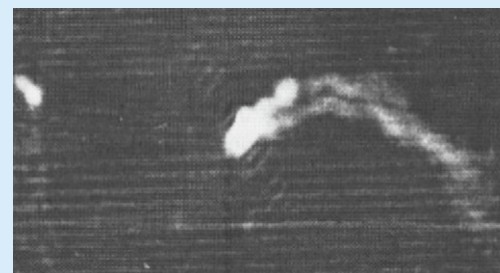
1. In the 1960s the only device that we could harness was a 'plotter', capable of drawing lines of arbitrary shape on a flat piece of paper. With it we projected mountainscapes out of parallel cross-sections of digital rasters. Our eye is reasonably sensitive to subtle elevation differences in this representation, but no more than that. And, of course, we may need additional view angles to see what is behind the mountains.

To use the plotter, one had to submit a 'job' to the University's mainframe computer.

*An early WSRT picture of M51. Its spiral arms are clearly visible and details of their position allowed speculation about their relation to the optical spirals.*



2. It would be much better if we could translate digital intensities photographically into grey scale. In principle one can replace the plotter paper by a photographic material and the pen by a light point, digitally controlling its brightness. The technique works, but is beset with many complications in practice.



*The 'photograph' that revealed the true nature of 3C129: A 'head-tail' source. The sensitive film is mounted on a rotating drum, the light point travels from left to right. The quasi-horizontal streaks is an artefact that is typical for mechanical arrangements with moving parts.*

3. Could we have the plotter draw dots, varying their surface density to represent grey shades? It turned out that plotter pens do not survive such abuse for more than a few thousand dots, but we did find a viable alternative: Replace the pen by a metal stylus striking against carbon paper in the way a classic typewriter does.

With a 5x5 dot raster per data point we can represent 26 shades of grey. In doing so we must correct for the strongly nonlinear relation between dot density and visual impression. Moreover, our eyes are experts in picking up unintended accidental patterns over large distances. The only way to suppress these was by adding random noise.

This plot technique was intensively used at ASTRON. Plotting a WSRT image took a whole night, but the result was worth waiting for.



*A plot made with the 5x5 dot-raster technique. (In successive Xerox, reproductions the crispness of the original image got lost).*

4. Around 1980 a plotter based on the Xerox technique appeared that could translate digital data directly into halftones. This precursor of modern printers would deliver an image in a few minutes. It was a great step forward, although it required frequent repair and maintenance and the printed images were often defaced by bands and stripes of deviant intensity.

At the same time colour monitors became available on which digital brightness and colour distributions could be faithfully and flexibly rendered. Only the Kapteyn Lab in Groningen managed to set aside the 70k guilders to buy one. For years on end it was used day and night and one had to reserve time in advance.

With the 1990 workstations the utopia from the 1960s became a reality. They marked the end of three decades of fascinating experiments on – or just beyond - the brink of the possible.



Ger de Bruyn discussing matters with Rick Perley from the NRAO. Ger has played a major role in the development of better data reduction software, as a motivator and as a high-profile “customer”. New techniques are accepted faster if they are espoused by a respected astronomer. Ger was prepared to work through the night to test new features, and never moaned but provided useful feedback. And he was always prepared to show and discuss some new puzzle on his screen.



The RSC software package contained some new features, like the use of parameterized source components as a sky model (rather than a pixellated image). This allowed a much more accurate subtraction of bright sources, leading to better images. After 1983, the RSC package was reimplemented by Wim Brouw, again in close collaboration with Ger de Bruyn. The resulting NEWSTAR package could be distributed to WSRT users to process the data themselves, but few took the trouble to learn yet another user interface (see also AIPS++ below).

The features of NEWSTAR reflect the pride of the WSRT group in the excellence of the WSRT, and the somewhat unusual determination to get the most out of the instrument. Its Selfcal solution could be made more robust by using the redundant spacing information as a powerful extra constraint that did not depend on a sky model. In addition, it could solve for many different (groups of) parameters, even for complex ones. It offered a new way to specify a subset of data to be processed, and was the first package to pioneer automatic flagging of “bad” data. All this prepared the way for generalized Selfcal (see MeqTrees below). Its polarization calibration created the conditions for discovering the Measurement Equation (see below). And it had a calculator (NCALC) that allowed users to play with the data.

The WSRT images produced by NEWSTAR have held the “Blue Ribbon” of Dynamic Range for three decades, until this distinction was taken over by MeqTrees. After Wim Brouw moved to Australia in 1991, he continued to help maintain NEWSTAR remotely, in exchange for a weekly box of Dutch cigars

(this is still the best value for money in the history of ASTRON). The package can still be used, and is even available online (in Github), but it effectively died with its only, but very worthy user, Ger de Bruyn, in 2017.

### The AIPS++ project

After 1980, there was a proliferation of software packages that exploited the Selfcal revolution. Since they all offered roughly the same functionality, many WSRT users reduced their data with their favourite package, notably AIPS or MIRIAD, rather than learning yet another user interface. In order to reach more users, it was tried for a while to implement some of the special features of NEWSTAR in the widely used AIPS package, but these mysteriously disappeared in the next release.

Still, it was widely recognised that the radio astronomy community would be served better by a single software package that was designed to deal with data from all the different telescopes, and maintained in such a way that new algorithms would quickly reach all users. The idea was adopted by Govert Croes, the new Director of Computing of NRAO, with the enthusiastic support of ASTRON. The result was the formation in 1991 of the AIPS++ consortium, by seven major radio astronomy institutes worldwide. (ATNF, ASTRON, BIMA, DRAO, MRAO, NRAO, TIFR).



The AIPS++ Steering Committee in 1994, and 25 years later. It consisted of self-propelled developers from the various institutes, who had all written their own software packages. Convinced of their own creations, “culture” tensions led the inevitable dynamics.

The outcome was a “qualified success”, i.e. it did not achieve its main goal. Eventually, the consortium did produce yet another package, now named CASA, which is primarily aimed at the processing data of the ALMA telescope and the upgraded VLA.

Apart from increasing the interaction between software developers worldwide, AIPS++ did produce some valuable by-products. Most notably the Measure-

ment Set (MS), a well-supported and generally accepted vehicle for the storage and access of visibility data. It was written by ASTRON's Ger van Diepen, who still maintains it as part of the "CasaCore" collection of software utilities. Other valuable modules include Measures and Fitting, both written by Wim Brouw. There was also talk about making it easy for users to contribute new software, in a so called Freedom Layer. And finally, the creative tension around AIPS++ served as a catalyst for the formulation of the Measurement Equation (see below) during the six months that ASTRON hosted the AIPS++ Calibration Design Team in 1992.

In retrospect, AIPS++ was doomed to fail because it would soon have turned into a White Elephant (see below). Nevertheless, it could have been much more successful for a while if we had started off by offering the users a vastly better package for handling uv-data than they were used to. This would have lured them to use AIPS++ for inspection and flagging, before turning to AIPS, MIRIAD or NEWSTAR for calibration and imaging. Only after they were hooked in this way should we have offered our own well-maintained set of well-tested applications, under a nifty user interface. Unfortunately, rather than letting the project grow by the pressure of customer enthusiasm, we started at the wrong end, turning potential users away.

So, even though the AIPS++ project did not achieve its ambitious goal, it was a necessary learning experience, with lots of useful lessons that might be heeded in a second attempt to address what is after all a real problem.

### *Formulation of the Measurement Equation (1995)*

Radio waves are transverse electromagnetic waves, and may be polarized. For a particular source, the polarized state is fully described by four Stokes parameters (I,Q,U,V). A full measurement of the incoming radiation requires two perpendicular dipoles in the focus of a dish, called X and Y in the case of the WSRT. A visibility sample then consists of four complex numbers, which are the result of correlating the two dipoles of the first dish with those of the second one, in all four combinations (XX,XY,YX,YY).

It took a remarkably long time to find a mathematical expression that describes a radio interferometer in full polarization. For more than 50 years, people used separate expressions for the four correlations, which contained approximations to simplify things. Until, in 1994, Johan Hamaker realised that the Kronecker matrix product would lead to an elegant expression without approximations. The instrumental behaviour of each antenna can be expressed as a multiplication of a sequence of 2x2 Jones matrices, each describing a separate effect on the incoming signal. Kronecker multiplication of the two antenna matrices then gives a 4x4 matrix that fully describes the conversion of the four Stokes parameters of the incoming radiation into the four correlations of a visibility sample.

This is now called the Measurement Equation (ME) of a generic radio interferometer. It is believed to describe interferometers with all kinds of antennas, which may not have to be of the same kind. A single dish can be described as a zero-baseline interferometer of the dish with itself.

The ME has been published in four elegant papers by Hamaker, Bregman and Sault in various combinations. It is also sometimes called the HBS formalism. The ME has later been extended for Direction-Dependent Effects (DDE) by Oleg Smirnov. MeqTrees is the only software package that fully implements the matrix formalism.



*Oleg Smirnov is standing on the shoulders of Johan Hamaker, Jaap Bregman and Bob Sault, the authors of the HBS-formalism, a.k.a. the Measurement Equation.*

### *The MeqTrees data reduction package*

The MeqTrees software package (Noordam, Smirnov) was started in 1995, partly out of exasperation with the AIPS++ project (see above), in which ASTRON was heavily involved. One of the bones of contention was the existence of Direction-Dependent Effects (DDE), like differences in antenna beamshapes. Such effects had been ignored in earlier packages, but were felt (by ASTRON) to be important in the much more sensitive new telescopes.

MeqTrees was designed to implement an arbitrary Measurement Equation (ME), with an arbitrary number of parameters. Moreover, it supports "generalized" Selfcal, in the sense that ME parameters may be smooth functions of time or frequency, or even sky position. The coefficients of such smooth functions (like low-order polynomials) then become the parameters of the ME. MeqTrees also pioneers new ways of visualizing the quality of the result, and of what is going on during the calibration process.

Obviously, there is a price to pay for solving for (substantially) more parameters. It takes more processing, and touches on deep questions of how many independent data points are needed to solve reliably for a given subset of parameters.



In the meantime, very fast algorithms have been developed to solve for the much smaller number of parameters of the simple traditional ME. For instance “Stef-Cal” which was developed in the SKADS context by the Oxford group, in collaborations with ASTRON’s Stefan Wijnholds. Because of the flexibility of MeqTrees, it has been the first to offer such techniques to the user. (Needless to say, users have not really taken up the offer, but have clung to their old packages). Nevertheless, MeqTrees has taken over the world record of Dynamic Range from NEWSTAR, first with WSRT observations, and now with data taken with the new VLA correlator (Smirnov, Perley). The software has moved to South Africa with Oleg Smirnov.

### Interaction with the world-wide software community

Prior to 1980, software developers were not really encouraged to visit sister institutes and interact with others who were trying to solve essentially the same problems. At ASTRON, it was considered sufficient for Wim Brouw to bring back enthusiastic reports of what he had learned abroad. This gradually changed with the UK/NL collaboration, which involved more travel for everyone, and with the advent of email and internet. The communication was greatly facilitated by the fact that the Dutch take easily to English, and had been exposed to the many foreigners that were attracted to the wonderful WSRT. After attempts to distribute NEWSTAR software as part of the much more widely used AIPS package had failed, the first real international software collaboration was the AIPS++ project (see above), which brought together the leading developers of seven major radio astronomy institutes. In the discussions, ASTRON insisted (unsuccessfully) on consistency in nomenclature and units, and tried to set right some sloppy practices that had crept in over the years.

ASTRON developers also started to attend ADASS, a yearly gathering of software developers in all branches of astronomy. Some even rose to membership of the Program Organizing Committee, and conducted a special discussion session on the “Future of Astronomical Data analysis Systems” (FADS). The latter were very popular (developers love grand-standing), and even came up with some ideas that might now be implemented with technology that did not exist in the 1990’s.

After the radio astronomy community became serious about building SKA and its precursors (e.g. LOFAR), ASTRON actively participated in yearly meetings about calibration and imaging issues (CALIM). As usual, ASTRON attempted (and failed) to impose a little structure on these discussions, in the hope of identifying priorities and perhaps steering some of the worldwide developments.

Finally, from 2000 onwards, it was felt that the tools now existed to collaborate effectively with a widely distributed group of developers. The primary motivation was to involve talented developers without luring them away from their home institutes. Apart from all kinds of internet tools, we expected great

things from PURR, a tool developed by Oleg Smirnov that made it really easy to report in great detail on a data reduction process. The MeqTrees group experimented with a “3GC Community”, in which developers worldwide could collaborate remotely in furthering 3rd Generation Calibration. Part of the plan was to offer simulated data-sets, which would reveal hidden delights to those that managed to reduce them well enough. Contenders might use any reduction package they liked, but it was assumed that MeqTrees would win.



From the above it might be concluded that all this interaction did not achieve anything tangible. That would be the wrong conclusion.

*Every year, experts from the worldwide SKA community gather at one of the participating institutes to discuss calibration and imaging issues (CALIM) pertaining to the new generation of giant radio telescopes. This one was in Dwingeloo in 2010.*

### A Tale of White Elephants

After half a century of processing the data taken with radio aperture synthesis telescopes, this might be a good moment to take stock of the situation. First of all, we can be proud of a phenomenal worldwide success story, in which ASTRON played a leading role throughout. In 5 decades, the quality of the result has improved by orders of magnitude in all important aspects. For instance, a widely used quality criterion like the Dynamic Range of our images has increased from 1:100 to 1:8.000.000 (eight million!)

We may also note with satisfaction that more sophisticated techniques and faster algorithms are still being developed apace, all over the world, and certainly also at ASTRON. Together with the rapid increase in speed of the hardware for computing, I/O and data transport, we will probably be able to process the huge data volumes from the new telescopes with sufficient accuracy. At the same time we must conclude that we have had, and still have, a substantial (worldwide) user interface problem.

The WSRT was the first aperture synthesis telescope with an open user policy, but felt compelled to reduce the unfamiliar visibility data for the users as a ser-



vice. This was unsatisfactory for more ambitious users, who wanted more control. After the Selfcal revolution of 1980, various self-propelled developers at different institutes created software packages like AIPS, NEWSTAR, MIRIAD, DIFMAP, etc, with which astronomers could reduce the data of particular telescopes themselves. They often did this in close collaboration with an ambitious user (Golden Teams), and were reasonably responsive to evolving needs. However, the packages all had their own user interfaces and data formats, and suffered from personality quirks and installation issues. Also, the ageing heroes found it increasingly difficult to adopt the newest developments. Still, single-hero packages evolve better than software under the care of teams, because the latter represent the priorities of institutes, and tend to have little emphatic contact with active users.

The result is that our valuable radio telescopes could be utilized much better by many more astronomers if the data reduction software would be easier to use. This includes looking for the hidden treasures in the very substantial data archives that have been accumulated over the years. After all, the large fields-of-view of radio telescopes contain much more than just the target object.

WIKIPEDIA defines a White Elephant as “a possession which its owner cannot dispose of and whose cost, particularly that of maintenance, is out of proportion to its usefulness. In modern usage, it is an object, building project, scheme, business venture, facility, etc., considered expensive but without use or value”

As it is, only users with special skills and/or resources (like teams) are able to cherry-pick the best features from the available software, to get results of the advertised quality. At the same time, existing software packages are turning into White Elephants, either because their creator dies, or they are under the care of institutional teams.

The AIPS++ project was an attempt during the 1990's to address the problem. A consortium of 7 leading institutes would jointly create a single package that would offer the best features of existing packages under a single user interface. It would be implemented with the latest technology (C++, OOP), would be well-maintained by an international team, and would smoothly evolve with the newest algorithms and technology. The story of its “qualified success” is a rich treasure trove of lessons that might be heeded in a next attempt. For the dream behind AIPS++ is still as valid as ever.

The key to a solution of this increasingly urgent problem is to recognise the proper role of, and relationship between, the various stake-holders, i.e. the telescope institutes, the self-propelled developers, and the scientific users. We should also look carefully at some highly successful models outside our small world, but this is outside the scope of a WSRT 50-year history.