

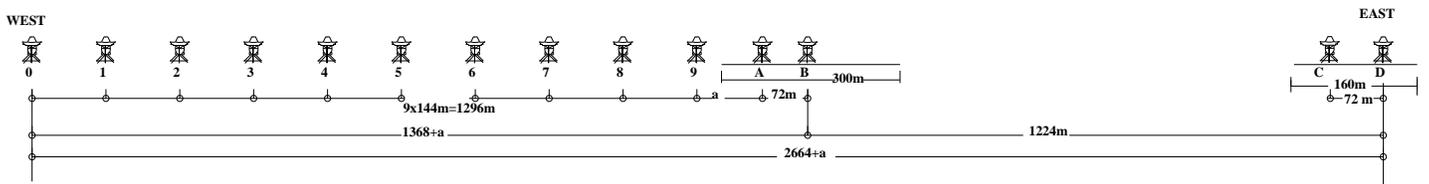
# ARSENAL USER DOCUMENTATION

## PART II

# INTRODUCTION TO THE THEORY OF APERTURE SYNTHESIS

July 16, 1993

Version: 1.0.0





# CONTENTS OF PART II

*This is part II of the WSRT User Documentation.*

*The WSRT User Documentation consists of 7 Parts; we may refer to other parts of the User Documentation. To get more information about how to obtain a copy and/or updates of parts of the User Documentation please send an e-mail request to [wsrc@nfra.nl](mailto:wsrc@nfra.nl)*

*The roman number in the pagenumber indicates the part.*

<b>1</b>	<b>Introduction</b>	<b>II-1-1</b>
<b>2</b>	<b>The measurement of visibilities</b>	<b>II-2-1</b>
2.1	The basic equation of aperture synthesis . . . . .	II-2-1
2.2	A simple correlating interferometer . . . . .	II-2-1
2.2.1	Response to a monochromatic point source . . . . .	II-2-2
2.2.2	Response to a monochromatic extended source . . . . .	II-2-2
2.3	Delay tracking . . . . .	II-2-4
2.3.1	The effect of bandwidth . . . . .	II-2-4
2.3.2	Delay tracking . . . . .	II-2-5
2.3.3	Bandwidth smearing . . . . .	II-2-6
2.4	Summary . . . . .	II-2-6
2.5	References . . . . .	II-2-7
<b>3</b>	<b>Recovering the brightness distribution</b>	<b>II-3-1</b>
3.1	Image formation and field of view . . . . .	II-3-1
<b>4</b>	<b>The Fourier transformation</b>	<b>II-4-1</b>
4.1	Introduction . . . . .	II-4-1
4.2	The Fourier transform . . . . .	II-4-1
4.3	The image and spectral domain . . . . .	II-4-2
4.4	Some important Fourier relations . . . . .	II-4-2
4.5	References . . . . .	II-4-6
<b>5</b>	<b>Fast Fourier transform effects</b>	<b>II-5-1</b>
5.1	FFT effects and functional relationships . . . . .	II-5-1
5.2	Synthesized beam and grading . . . . .	II-5-5
5.3	Grating Responses . . . . .	II-5-7
5.4	Errors, Sidelobes and Confusion . . . . .	II-5-7
5.5	References . . . . .	II-5-9
<b>6</b>	<b>The coordinate system</b>	<b>II-6-1</b>
6.1	$u, v$ and $l, m$ coordinates . . . . .	II-6-1
6.2	The difference between 3d and E-W coordinates . . . . .	II-6-3
6.3	References . . . . .	II-6-3

---

<b>7</b>	<b>Temperature, brightness and sensitivity</b>	<b>II-7-1</b>
7.1	Antenna temperature . . . . .	II-7-1
7.2	Brightness Temperature . . . . .	II-7-2
7.3	Relations between Flux density, Brightness and Temperatures . . . . .	II-7-3
7.3.1	Estimates of $\Omega_A$ . . . . .	II-7-3
7.3.2	Conversion factor $T_b(\text{K})-S(\text{mJy})$ . . . . .	II-7-4
7.4	Sensitivity . . . . .	II-7-4
<b>8</b>	<b>Effects in Fourier transformed spectra</b>	<b>II-8-1</b>
8.1	Introduction . . . . .	II-8-1
8.2	Phase effects . . . . .	II-8-6
8.3	Tapering . . . . .	II-8-9
8.4	Delay effects . . . . .	II-8-10
8.5	Fast Fourier Transform effects . . . . .	II-8-10
8.6	Noise Distribution in Tapered Spectra . . . . .	II-8-11
8.7	Other tapers . . . . .	II-8-13
8.8	References . . . . .	II-8-13
<b>9</b>	<b>Bibliography and References</b>	<b>II-9-1</b>
9.1	Book reviews . . . . .	II-9-1
9.2	Dictionary . . . . .	II-9-5

# WSRT PART II

## USER DOCUMENTATION CHAPTER 1

## INTRODUCTION

In this part we present some of the basic theory of aperture synthesis and related techniques. Our intention is to present a text which can be used by astronomers who need to refresh their basic knowledge of aperture synthesis. Scientists for whom this text is the first encounter with radio interferometry are advised to read one of the books discussed in section 9.1.

The WSRT is an aperture synthesis telescope. It combines thousands of interferometer measurements to construct an image as if this image was measured by a large single aperture antenna. The quantity that is measured by an interferometer is called a visibility. From a set of visibility measurements one can obtain a brightness distribution using a Fourier transformation.

In chapter 2 we explain how an interferometer can be used to measure the visibility. In chapter 3 we explain how we obtain a brightness distribution from the visibilities. The Fourier transformation and its properties are discussed in chapter 4.

Although the Fourier transformation from visibility to brightness distribution seems rather straightforward there are a number of complications. One of those complications is for example that visibilities are measured as function of baseline separation and orientation. In order to get a complete and unambiguous description of the brightness distribution one needs to measure the visibility in all orientations and on all baseline lengths from zero out to infinity. This is of course not possible. Another complication is that the numerical implementation of the Fourier transformation (Fast Fourier transformation) interpolates the visibility data onto a computational grid. Both effects act as filters and introduce artifacts in the astronomical images. In chapter 5 we explain the effects of these and other “filters” and discuss the measures that are taken to reduce the unwanted side-effects. Our purpose is to obtain the fluxes or brightness temperatures as a function of position in RA,DEC on the sky, in chapter 6 we discuss the coordinate systems used in Westerbork. In chapter 7 we discuss the relation between temperatures and brightness. We also discuss how the sensitivity of a telescope can be calculated. We only discuss the theory, for the sensitivity of the WSRT see part III (“Specific Aspects of the WSRT Synthesis Telescope”), chapter 2.

The whole discussion in chapter 2, 3, and 5 applies to ‘monochromatic’ images. In practice more frequency bands (channels) are measured at the same time. The relations between these measurements and the frequency spectrum are roughly the same as the relations between visibility and brightness distribution. The Fourier relations between the measurement and the spectrum and the artifacts introduced when producing spectra are discussed in chapter 8.

The closing chapter (9) contains a bibliography and references to important articles. We also included a dictionary to help you cope with jargon.

# THE MEASUREMENT OF VISIBILITIES

*by Olaf Kolkman. Based on texts and pictures of the books discussed in section 9.1*

In this chapter we will define visibility and we will show how visibilities are measured with a simple interferometer. We keep the mathematical description as simple and as general as possible.

Notation: Because we deal with electromagnetic waves, a complex quantity, we use the notation  $z = |z|e^{j\Phi} = |z|(\cos(\Phi) + j\sin(\Phi))$  where  $|z|$  is the amplitude and  $\arg(z)=\Phi$  is the phase of the wave. We will denote the real and imaginary part of  $z$  by  $\Re(z) = |z|\cos(\Phi)$  and  $\Im(z) = |z|\sin(\Phi)$  respectively. Boldfaced letters like  $\mathbf{s}$  denote vectors.

## 2.1 THE BASIC EQUATION OF APERTURE SYNTHESIS

The basic relation of aperture synthesis can be written as:

$$\mathcal{V}(\mathbf{D}_\lambda) = \int_{4\pi} B'(\sigma) e^{-j2\pi\mathbf{D}_\lambda \cdot \sigma} d\Omega \quad (2.1)$$

where  $\mathcal{V}(\mathbf{D}_\lambda)$  is a complex function called the visibility function. It can be measured using a two element interferometer separated by the vector  $\mathbf{D}_\lambda$ , whose length is expressed in units of wavelength.  $\sigma$  is the difference between an arbitrary direction vectors and the pointing direction vector of the telescope  $\mathbf{s}_0$ , which also is referred to as the phase reference point or phase center. The quantity  $B'(\sigma)$  is directly related to the brightness distribution of the source, the quantity we are interested in. The visibility  $\mathcal{V}(\mathbf{D}_\lambda)$  can be measured using a correlating interferometer.

The functional form of equation 2.1 is that of a Fourier transform (chapter 4). Equation 2.1 can be inverted to obtain  $B'(\sigma)$  from which we can obtain the real brightness distribution  $B(\sigma)$ .

Note that we used a generalized coordinate system here. In order to invert equation 2.1 we have to choose a specific coordinate system. A discussion of the coordinate system will be given in chapter 6.

## 2.2 A SIMPLE CORRELATING INTERFEROMETER

In this and the following section we will illustrate how in theory visibilities can be obtained using an interferometer. In this treatment we use a simplified interferometer. A working interferometer is much more complicated. We refer to Thompson *et al.* (1986) and Perley *et al.* (1989) for a more detailed and complete discussion.

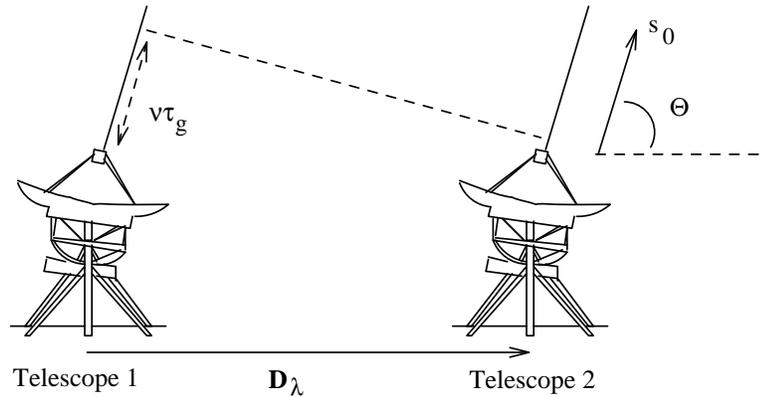


Figure 2.1: A simple interferometer

### 2.2.1 RESPONSE TO A MONOCHROMATIC POINT SOURCE

The response of a simple correlating interferometer to a point source radiating monochromatic radiation is a sinusoidal function whose frequency depends on the spacing between the antennas and the direction of viewing.

Consider two identical parabolic mirrors separated by a vector  $\mathbf{D}_\lambda$  (figure 2.1) measuring monochromatic radiation of frequency  $\nu$  from a point source in the direction  $\mathbf{s}_0$ .  $|\mathbf{D}_\lambda|$  is expressed in units of wavelength.  $\mathbf{s}_0$  is a vector of unit length. We will assume a plane parallel wavefront and ignore atmospheric and other disturbances.

The signal in telescope 2 will be proportional to  $\sin(2\pi\nu t)$ . The signal in telescope 1 will be phase shifted by an amount  $2\pi\nu\tau_g$ , where  $\tau_g$  is the geometrical delay. It can be shown that  $\nu\tau_g = \mathbf{D}_\lambda \cdot \mathbf{s}_0$ . The response of a multiplying correlator will be:

$$\begin{aligned} r &\propto 2 \sin(2\pi\nu t) \sin(2\pi\nu(t - \tau_g)) \\ &= \cos 2\pi\nu\tau_g - \cos(4\pi\nu t) \cos(2\pi\nu\tau_g) \\ &\quad - \sin(4\pi\nu t) \sin(2\pi\nu\tau_g) \end{aligned} \quad (2.2)$$

The sine and cosine terms with  $t$  are oscillating rapidly. These high frequency terms ( $\cos(4\pi\nu t) \cos(2\pi\nu\tau_g)$  and  $\sin(4\pi\nu t) \sin(2\pi\nu\tau_g)$ ) are filtered out leaving the low frequency term:

$$r \propto F(\mathbf{D}_\lambda, \mathbf{s}_0) = \cos 2\pi\nu\tau_g = \cos(2\pi\mathbf{D}_\lambda \cdot \mathbf{s}_0) \quad (2.3)$$

When tracing the point source, the direction of the vector  $\mathbf{s}_0$  varies with time so the output of the correlator will be a sinusoidal function  $F$  depending on  $\mathbf{D}_\lambda$  and  $\mathbf{s}_0$ . The function  $F$  is called the fringe function. The frequency of  $F$  is a function of  $|\mathbf{D}_\lambda|$  (figure 2.2).

### 2.2.2 RESPONSE TO A MONOCHROMATIC EXTENDED SOURCE

We have shown above what the response of an interferometer to a monochromatic point source is. We will now investigate what the response of a simple interferometer to an extended source is. We will define the visibility.

Let us now point our simple interferometer at an extended source with surface brightness  $B(\sigma)$ , radiating quasi monochromatically at a frequency  $\nu$  in a band of width  $d\nu$ .

The telescopes are pointed towards the center of the source in the direction  $\mathbf{s}_0$ ; this point is referred to as the phase reference point or the fringe stopping center. A sky area of size  $d\Omega$  in the direction  $\mathbf{s} = \mathbf{s}_0 + \sigma$  will contribute a component of power in each of the antennas that is proportional to  $A(\sigma)B(\sigma)d\nu$ , where  $A(\sigma)$  is the reception pattern or power pattern which describes the sensitivity of the antenna element for radiation coming

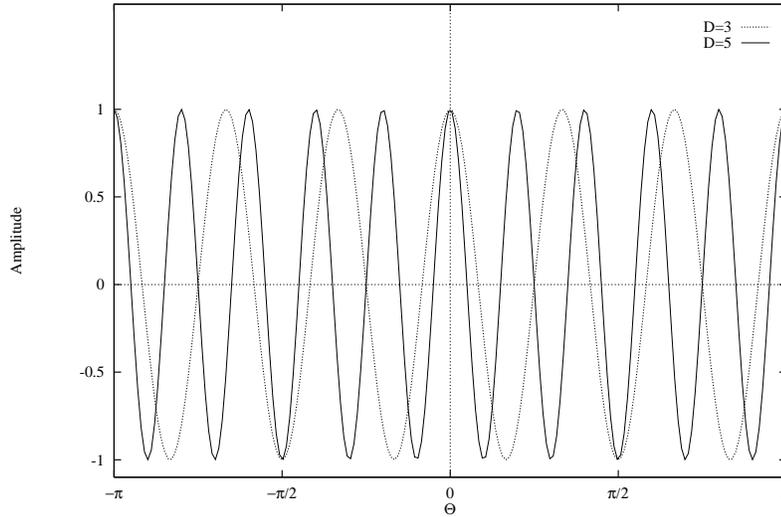


Figure 2.2: The fringe function for two different telescope spacings  $D_\lambda = 3$  and  $D_\lambda = 5$ . The angle  $\theta$  is the hour angle in radians. The amplitude is normalized.

from different directions. The output of the correlator due to radiation from direction  $\mathbf{s}$  will be proportional to the power in the antennas from direction  $\mathbf{s}$  and the fringe function in direction  $\mathbf{s}$ ,  $F(\mathbf{D}_\lambda, \mathbf{s}) = F(\mathbf{D}_\lambda, \mathbf{s}_0 + \sigma)$ . The response of the interferometer to radiation from the source can be obtained by integrating over the source, assuming that signals from two different directions do not correlate i.e. the source is spatially incoherent:

$$\begin{aligned}
 r(\mathbf{D}_\lambda, \mathbf{s}_0) &= d\nu A(0) \int_{4\pi} A_N(\sigma) B(\sigma) \cos[(2\pi \mathbf{D}_\lambda \cdot (\mathbf{s}_0 + \sigma))] d\Omega \\
 &= d\nu A(0) \cos(2\pi \mathbf{D}_\lambda \cdot \mathbf{s}_0) \int_{4\pi} A_N(\sigma) B(\sigma) \cos(2\pi \mathbf{D}_\lambda \cdot \sigma) d\Omega \\
 &\quad - d\nu A(0) \sin(2\pi \mathbf{D}_\lambda \cdot \mathbf{s}_0) \int_{4\pi} A_N(\sigma) B(\sigma) \sin(2\pi \mathbf{D}_\lambda \cdot \sigma) d\Omega
 \end{aligned} \tag{2.4}$$

where  $A_N(\sigma) = A(\sigma)/A(0)$  is the normalized antenna reception pattern. We remind you that  $r(\mathbf{D}_\lambda, \mathbf{s}_0)$  should be read as the response of a correlating interferometer with baseline vector  $\mathbf{D}_\lambda$  pointed at  $\mathbf{s}_0$ . We now *define* the visibility as

$$\mathcal{V} \equiv \int_{4\pi} B'(\sigma) e^{(2\pi j \mathbf{D}_\lambda \cdot \sigma)} d\Omega = |\mathcal{V}| e^{j \Phi_{\mathcal{V}}} \tag{2.5}$$

The phase of  $\mathcal{V}$ ,  $\Phi_{\mathcal{V}}$  is measured relative to the phase reference point at  $\mathbf{s}_0$ , and  $B'(\sigma) = A_N(\sigma)B(\sigma)$  is the modified brightness distribution. The real and imaginary parts of  $\mathcal{V}$  can be separated so we obtain:

$$\int_{4\pi} A_N(\sigma) B(\sigma) \cos(2\pi \mathbf{D}_\lambda \cdot \sigma) d\Omega = |\mathcal{V}| \cos \Phi_{\mathcal{V}} \tag{2.6}$$

$$\int_{4\pi} A_N(\sigma)B(\sigma) \sin(2\pi\mathbf{D}_\lambda \cdot \sigma) d\Omega = -|\mathcal{V}| \sin \Phi_\mathcal{V} \quad (2.7)$$

Substituting 2.6 an 2.7 in 2.4 we can rewrite the response as:

$$r(\mathbf{D}_\lambda, \mathbf{s}_0) = d\nu A(0)|\mathcal{V}| \cos(2\pi\mathbf{D}_\lambda \cdot \mathbf{s}_0 - \Phi_\mathcal{V}) \quad (2.8)$$

The equation above also demonstrates that the response to a monochromatic extended source, *i.e.* the output of the correlator, is a fringe pattern with a frequency corresponding to that of a hypothetical point source at the position  $\mathbf{s}_0$ . The amplitude of the fringes is proportional to the visibility and when phases are measured relative to the response of the source at  $\mathbf{s}_0$  then the phase of the response is that of the visibility.

If we take a close look at the equation above we can see that the response to a monochromatic extended source is a direct measure of the visibility as defined in equation 2.5. The amplitude of the fringes is proportional to the visibility amplitude. And the phase of the response, relative to the phase in direction  $\mathbf{s}_0$ , is the phase of the visibility,  $\Phi_\mathcal{V}$ .

Note that:

- We are using generalized coordinates. In practice a coordinate system will be used in which the visibility and the response is only a function of projected baselines  $(u, v)$  ( $\sqrt{u^2 + v^2} = \mathbf{D}_\lambda \cdot \mathbf{s}_0$ ).
- The response to a point source at  $\mathbf{s}$  is exactly the response of a point source at  $\mathbf{s}_0$ , but phase shifted by an amount proportional to the difference in path length between the two telescopes if pointed at  $\mathbf{s}$  instead of  $\mathbf{s}_0$ , *i.e.*  $\Delta\Phi = 2\pi(\mathbf{D}_\lambda \cdot \mathbf{s} - \mathbf{D}_\lambda \cdot \mathbf{s}_0)$ .

## 2.3 DELAY TRACKING

### 2.3.1 THE EFFECT OF BANDWIDTH

In practice we will not have an infinitesimally bandwidth. Let us first consider the effect on the fringe pattern if we are observing at two different frequencies,  $\nu_1$  and  $\nu_2$  separated by a frequency  $\Delta\nu$ . The two fringes will add incoherently to a pattern as shown in figure 2.3.

At a certain phase  $\Delta\Theta$  there will be destructive interference. The correlator response, *i.e.* the product of the visibility and the fringe function, will be zero at that phase and we can not measure anything at all. This sets, at a given bandwidth, a limit on the geometrical delay  $\tau_g$  for which there is correlation. Using equation 2.3 we can calculate that  $F(\tau_g) \geq 0$  for  $-\pi > 2\pi\nu\tau_g > \pi$  so fringes will be measurable only if  $\Delta\tau_g \leq \frac{1}{\Delta\nu}$ .

In general the fringe pattern has an envelope determined by the Fourier transform of the instrumental frequency response, usually called the fringe washing function. This is discussed in Thompson *et al.* (1986, section 2.2 page 44). As an example we will show the effect of a box shaped bandpass filter on the fringe pattern.

Let us for the moment consider the simple case of the two element interferometer with a receiving system that has uniform power response over a band of width  $\Delta\nu$  centered on  $\nu_0$  (the bandshape function  $H(\nu)$  is zero elsewhere). The antennas are pointed at a point source. The response to the point source in the infinitesimal band  $d\nu$  at frequency  $\nu$  is a product of the modified brightness distribution  $B(\mathbf{s}_0)$ , and the fringe function,  $F$ :

$$dr = d\nu B(\mathbf{s}_0) \cos(2\pi\nu\tau_g)$$

This has to be integrated over all frequencies:

$$\begin{aligned} r &= \frac{1}{\Delta\nu} \int_0^\infty d\nu B(\mathbf{s}_0) \cos(2\pi\nu\tau_g) H(\nu) \\ &= \frac{1}{\Delta\nu} \int_{\nu_0 - \Delta\nu/2}^{\nu_0 + \Delta\nu/2} d\nu B(\mathbf{s}_0) \cos(2\pi\nu\tau_g) \\ &= \cos(2\pi\nu_0\tau_g) \frac{\sin(\pi\Delta\nu\tau_g)}{\pi\Delta\nu\tau_g} B(\mathbf{s}_0) \end{aligned} \quad (2.9)$$

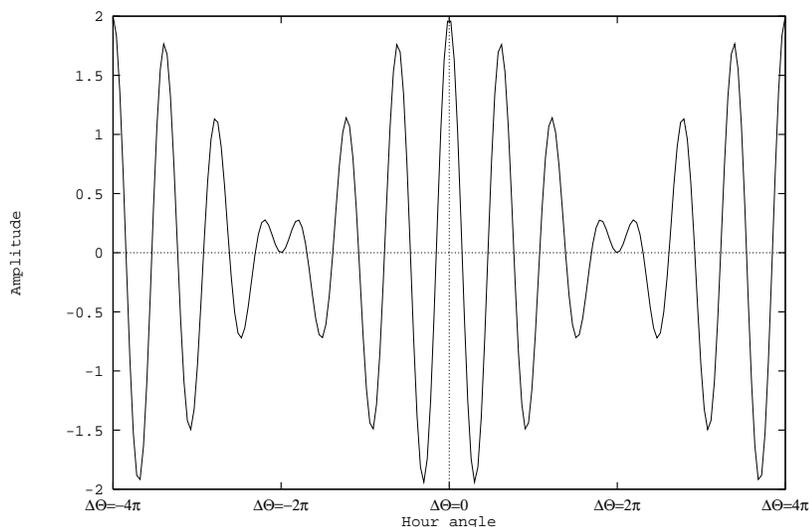


Figure 2.3: When two fringes of frequency  $\nu = 3$  and  $\nu = 3.5$  are added an interference pattern will result. At hour angle  $\theta = 2\pi \pm 2k\pi (k \in \mathcal{N})$  destructive interference will occur

The sine function of the above equation determines the envelope of the sinusoidal fringe. The response is shown in figure 2.4. Note that the amplitude of the response is only proportional to the visibility amplitude for  $\tau_g = 0$ .

### 2.3.2 DELAY TRACKING

We have seen that if we track a source using an interferometer with a finite bandwidth the amplitude of the response will be modulated by the fringe washing function *i.e.* the Fourier transform of the bandpass. To obtain the highest response possible we delay the signal of one of the telescopes by an amount  $\tau_i \approx \tau_g (= \frac{1}{\nu} \mathbf{D}_\lambda \cdot \mathbf{s}_0)$ . The effect of delay tracking when observing a point source at  $\mathbf{s}_0$  is that the path length difference is reduced to zero so the fringe term will be unity and we are measuring a constant signal proportional to the amplitude of the point source.

The effect of delay tracking when measuring an extended source is more complex and can be understood from inspecting equation 2.8 for a delay tracking telescope:

$$r(\mathbf{D}_\lambda, \mathbf{s}_0) = d\nu A(0) |\mathcal{V}| \cos(2\pi \mathbf{D}_\lambda \cdot \mathbf{s}_0 - 2\pi \nu \tau_i - \Phi_\nu)$$

Integrating over frequency, remembering that the fringe term  $\mathbf{D}_\lambda \cdot \mathbf{s}_0 - \nu \tau_i$  is set to 0 and that the cosine function is symmetric we obtain:

$$r(\mathbf{D}_\lambda, \mathbf{s}_0) = CA(0) |\mathcal{V}| \cos \Phi_\nu \quad (2.10)$$

where C is an integration constant. From the above we see that using a delay tracking interferometer we obtain  $|\mathcal{V}|$  and  $\Phi_\nu$ .

At Westerbork a digital delay is implemented which corrects the signal at video frequencies. The delay is implemented such that the delay is constantly changed as the hour angle of the pointing center changes. If the delay would not be continuously changed the fringes would oscillate with the natural fringe frequency (also see section 10.2 of part III). The process of reducing the fringe frequency to zero by maintaining  $2\pi(\mathbf{D}_\lambda \cdot \mathbf{s}_0 - \nu \tau_i) = 0$  is called fringe stopping or fringe rotation.

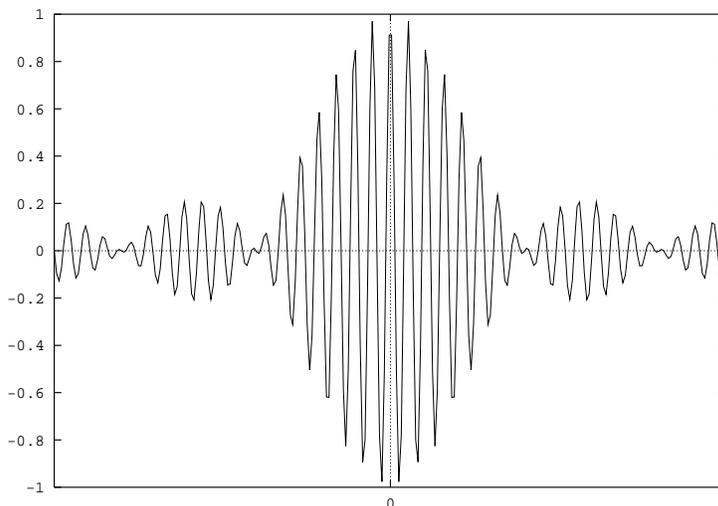


Figure 2.4: The point-source response of an interferometer with a box shaped frequency passband. The abscissa is the geometrical delay  $\tau_g$ .

### 2.3.3 BANDWIDTH SMEARING

Another effect of using a finite bandwidth is bandwidth smearing. Bandwidth smearing distorts the edges of large broadband (continuum) maps. This is due to the fact that different frequencies have different baseline lengths (in units of wavelength). Fourier transforming the visibilities assuming a baseline length equal to the geometrical baseline length divided by the central wavelength will cause an underestimation of the baseline length at lower frequencies and overestimation at the higher frequencies. This will introduce phase errors which will lead to image distortions.

A mathematical treatment of these matters is given in chapter 13 in the textbook by Bridle and Schwab in Perley *et al.* (1989).

Bandwidth smearing is a form of chromatic aberration.

## 2.4 SUMMARY

In the above we have shown that we can measure the phase and amplitude of the visibility of a source using a delay tracking interferometer. We have also shown that the visibility is a function of  $\mathbf{D}_\lambda$ ,  $B(\sigma)$ , and  $\mathbf{s}_0$  only. This means that the visibility can be expressed as a function of the projected baselines.

Remember that three assumptions are made:

1. The source is at such a distance that the incoming radio rays are parallel to the antennas. This is a valid assumption for most interferometers except for VLBI or millimeter interferometers observing objects in the solar system.
2. The atmosphere does not influence the phase of the incoming waves and the equipment is stable. This is of course never the case. See chapter 3 in part IV for more details.
3. Radiation from two different places in the source does not correlate *i.e.* the radiation is spatially incoherent. This is a very important assumption. When sources are not spatially incoherent the relation between brightness distribution and visibility is not a simple Fourier transform any more. (e.g. chapter 3

in Thompson *et al.*(1986) or equation (1-3) in chapter 1 by Clark in Perley *et al.*(1989)). The term spatial coherence function is sometimes used by authors as the quantity that is measured by interferometers.

The coherence function,  $\Gamma$ , is a function of spatial frequencies  $u$ ,  $v$  and time  $\tau$ . The real part of this function is basically what is measured by the WSRT digital correlator.  $\Gamma(u, v, \tau)$  relates to the visibility as:

$$\mathcal{V}(u, v, \nu) = \int_{-\infty}^{\infty} \Gamma(u, v, \tau) e^{-2\pi\nu\tau} d\tau \quad (2.11)$$

where  $\nu$  is the frequency of the observation.

Equation 2.11 is a temporal Fourier transform. It is of importance when discussing line observations. For more details see the books mentioned below.

For a discussion of aperture synthesis in terms of the coherence function one is also referred to the books mentioned below.

The system described above is an ideal simplified interferometer. Keep in mind though that in a working interferometer the signal is being processed by amplifiers, filters, mixers, etc. each having their own gain and introducing phase shifts to the signal.

## 2.5 REFERENCES

- Perley, R.A., Schwab, F. and Bridle, A.H. (Editors) (1989): "*Synthesis Imaging in Radio Astronomy*". Astronomical Society of the Pacific.. ISBN:0-937707-23-6. (see also review in part II Chapter 9.1, Book 2).
- Thompson, A.R., Moran, J.M. and Swenson Jr., G.W. (1986): "*Interferometry and Synthesis in Radio Astronomy*". John Wiley & Sons, New York.. ISBN 0-471-80614-5. (see also review in part II Chapter 9.1, book 3).

# RECOVERING THE BRIGHTNESS DISTRIBUTION

*Based on a text from the old WSRT manual that was edited by A.G. Willis and originally based on articles by J.A.Högbom, R.H. Harten and W.N. Brouw.*

In chapter 2 we have seen that an interferometer measures complex visibilities  $\mathcal{V}(u, v)$  and that the visibility function can be transformed to obtain the brightness distribution.

## 3.1 IMAGE FORMATION AND FIELD OF VIEW

Let us rewrite equation 2.1 in the appropriate coordinates <sup>1</sup>

$$\mathcal{V}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(l, m, \delta_0) A_N(l, m) B(l, m) e^{-j2\pi(u l + v m)} dl dm \quad (3.1)$$

The function  $f(l, m, \delta_0)$  contains some additional terms (e.g. deviations from the ideal interferometer behavior) which must be corrected for in the actual data reduction but need not be considered here further.

We may invert equation 3.1, by means of an inverse Fourier transform, to obtain

$$B(l, m) A_N(l, m) = \frac{1}{f(l, m, \delta_0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{V}(u, v) e^{-2\pi j(u l + v m)} du dv \quad (3.2)$$

In practice we are not able to measure at all spacings, or  $(u, v)$  points, out to infinity so let us see how the data is sampled.

Because  $B(l, m)$  and  $A_N B(l, m)$  are real functions of  $(l, m)$  it can be shown from the properties of the Fourier transform that

$$\mathcal{V}(-u, -v) = \mathcal{V}^*(u, v) \quad (3.3)$$

Thus it is sufficient to measure the complex visibility  $\mathcal{V}(u, v)$  over two adjacent quadrants in the  $(u, v)$  plane since the data in the other half of the plane can be determined from equation 3.3.

---

<sup>1</sup>See chapter 6 for a discussion of the coordinate system which is used here. To understand this discussion it is enough to know that the  $l, m$  coordinates are used to describe the brightness distribution and the  $u, v$  coordinates are used to describe the visibility distribution

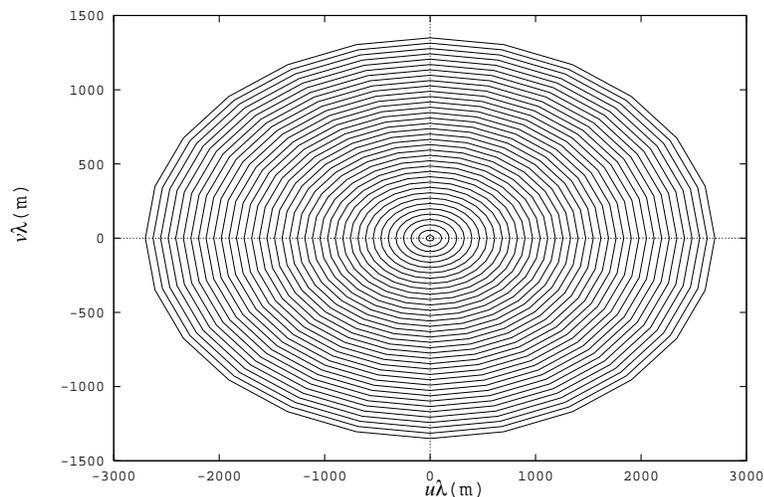


Figure 3.1: If we observe a source at a declination of  $30^\circ$ , with a standard line-setup, for a period of 12 hours, we will cover the  $(u, v)$  plane with 38 ellipses (or 40 depending on the position of the movable telescopes).

For an E-W interferometer we can rewrite the  $u, v$  coordinates as a function of baseline length and source position in declination and hour angle (cf. 6.3 and 6.4 in chapter 6)

$$\begin{aligned} u &= D_\lambda \cos h_0 \\ v &= D_\lambda \sin h_0 \sin \delta_0 \end{aligned}$$

It can be seen from these equations that when an interferometer tracks a source over a range of hour-angles, the projected interferometer baseline will trace out an ellipse in the  $(u, v)$  plane. In a 12 hour measurement, during which the earth rotates the interferometer baselines over  $180^\circ$ , all quadrants in the  $(u, v)$  plane will be covered with ellipses because of the symmetry of the visibility function ( $\mathcal{V}(-u, -v) = \mathcal{V}^*(u, v)$ ). In a 12 hour measurement the WSRT (in standard line setup) samples 40 ellipses<sup>2</sup> in the  $(u, v)$  plane. The ellipticity depends on the declination of the source. The size of the major axis of a certain ellipse depends on the distance between two particular telescopes. An example is shown in figure 3.1. When a redundant configuration is used up to 91 ellipses, of which about half will overlap with others, can be produced.

<sup>2</sup>N.B. actually 38 baselines are sampled most of the time, as 0A=9C and 0B=9D

# THE FOURIER TRANSFORMATION

*Based on a text from the old WSRT manual. The text was originally edited by A.G. Willis based on articles by J.A.Hógbom, W.N.Brouw and R.H.Harten. The text has been updated and the original figures have been redrawn by O.M. Kolkman*

## 4.1 INTRODUCTION

In the previous chapters we have shown that we can, in theory, obtain a brightness distribution from visibility data using Fourier transforms. In this chapter we review the basic theory of those transformations.

## 4.2 THE FOURIER TRANSFORM

The Fourier transform of the function  $f(x)$  is defined as

$$F(s) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi xs} dx \quad (4.1)$$

The inverse transformation is defined as:

$$f(x) = \int_{-\infty}^{\infty} F(s)e^{j2\pi xs} ds \quad (4.2)$$

In the two dimensional case the definitions are similar. The Fourier transform and its inverse are written as

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)e^{-j2\pi(xu+yv)} dx dy \quad (4.3)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v)e^{j2\pi(xu+yv)} dv du \quad (4.4)$$

The properties of the Fourier transform are described in several text books among which the book by Bracewell (1978) is a standard work. Several properties relevant to the subject of image processing are described in

Gonzalez and Wintz (1987). This book also has a description of the fast Fourier transform (FFT) which is the algorithm used in most of the computer applications.

Fourier transformation can be conceived more easily if we rewrite equation 4.2 in a discrete form:

$$f(x_i) = \sum_{k=0}^{N-1} F(u_k) e^{j2\pi u_k x_i / N} \quad (4.5)$$

for  $x_i = i\Delta x$  for  $i = 0, 1, 2, \dots, N-1$  and  $u_k = k\Delta u$  for  $u = 0, 1, 2, \dots, N-1$

We see that the function  $f(x)$  is the sum of complex oscillators ( $F(u)e^{j2\pi ux/N}$ ) of frequency  $u$ , with an amplitude  $|F(u)|$  and phase shifted by an amount  $\arg(F(u))$ . The Fourier transform projects the function  $f(x)$  on  $F(u)$  i.e. decomposes the function  $f(x)$  in oscillators.

The amplitude  $|F(u)|$  is often called the Fourier spectrum of  $f(x)$ . It describes the relative contribution of an oscillator with frequency  $u$  in  $f(x)$ .

### 4.3 THE IMAGE AND SPECTRAL DOMAIN

In radio astronomy one often talks about the image or spatial domain and spatial frequency or  $u, v$ -domain. In the image domain information is available in the way most people conceive information of objects i.e. as a brightness distribution as a function of  $x$  and  $y$ . The spectral domain is a more abstract domain. It both contains the relative amplitudes as well as the phases of the oscillators describing structures with spatial frequencies  $(u, v)$

The Fourier transform transforms objects in the first domain into the second domain. A large structure in the spatial domain will give rise to power at a point in the  $(u, v)$ -plane near the origin. The reciprocal of  $\sqrt{u^2 + v^2}$  is a measure of the size of structure and the amplitude of the complex visibility point is a measure of power. Note that  $u$  and  $v$  are real but  $\mathcal{V}(u, v)$  is a complex number. Small structures in the brightness distribution will give rise to structure extending far from the origin of the  $(u, v)$ -plane.

Using synthesis techniques the visibilities are measured. As stated in equation 2.1 the brightness distribution is the Fourier transform of the visibility. The visibility is a function in the spectral domain and contains information about the spatial frequencies in the measured object.

Some functions  $f(x, y)$  and their Fourier Transform  $F(u, v)$  are shown in figure 4.1. They may be interpreted as brightness distributions and their associated visibilities.

### 4.4 SOME IMPORTANT FOURIER RELATIONS

Some of the important theorems about the properties of Fourier transforms are summarized below. The reader is referred to Bracewell (1978), Chapter 6, for derivations and proofs.

**SIMILARITY THEOREM** Scaling of the argument of a function will scale the Fourier transform of that function and the argument of that Fourier transform.

If  $f(x) \rightleftharpoons F(s)$  then  $f(ax) \rightleftharpoons |a|^{-1}F(s/a)$ .

**ADDITION THEOREM** (also called the linearity theorem)

The Fourier transform of a sum of functions equals the sum of the Fourier transforms of the individual functions.

If  $g(x) \rightleftharpoons G(s)$ ,  $f(x) \rightleftharpoons F(s)$ ,  $h(x) \rightleftharpoons H(s)$  and  $h(x) = f(x) + g(x)$  then  $H(s) = F(s) + G(s)$ .

**TRANSFORMATION OF FUNCTIONS** Let us define even and odd (complex) functions. A function  $f_e(x)$  is even if  $f_e(-x) = f_e(x)$ . A function  $f_o(x)$  is odd if  $f_o(-x) = -f_o(x)$ . The Fourier transform of an even function is even. The Fourier transform of an odd function is odd. In general a complex function  $f(x)$  can always be

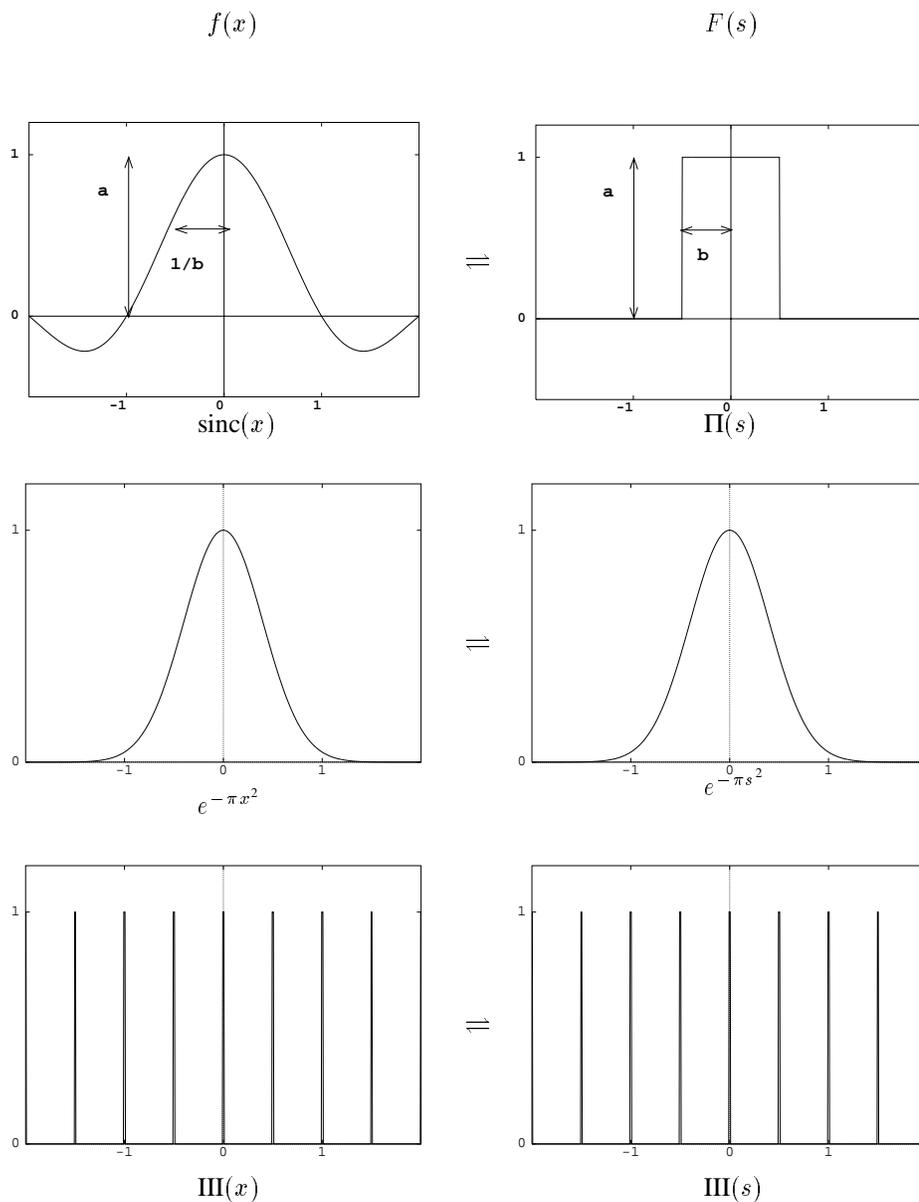


Figure 4.1: Some functions  $f(x)$  and their Fourier transforms  $F(s)$  are drawn. The  $x$  relates to  $s$  as  $s = \frac{1}{x}$ . The *shah* function is defined as  $\text{III}(X) = \sum_{n=-\infty}^{\infty} \delta(x - n)$ . The *shah* function is important when sampling data. The *sinc* function is defined as  $\text{sinc}x = \frac{\sin x}{x}$ . Its Fourier transform, the *unit rectangle function* is defined as  $\Pi = 1$  for  $|x| < \frac{1}{2}$  and  $\Pi = 0$  for  $|x| > \frac{1}{2}$ .

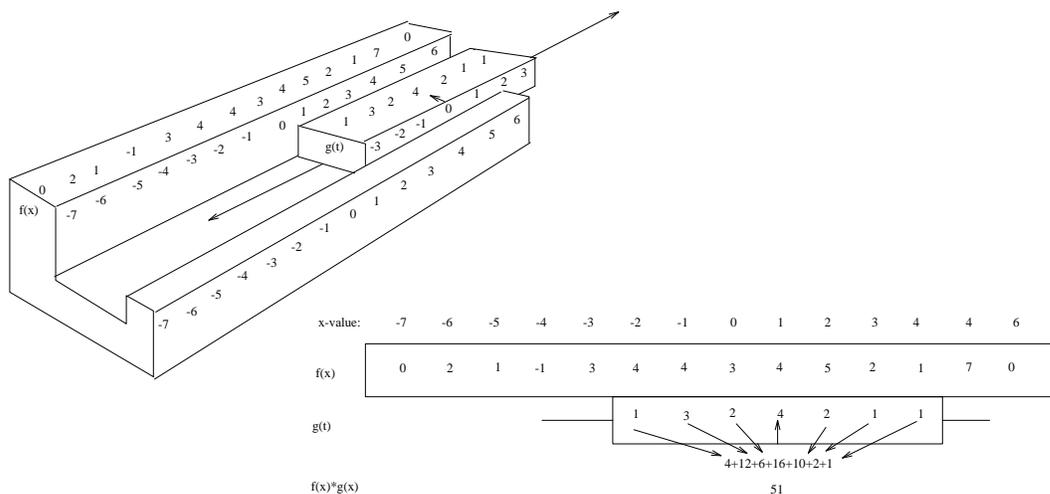


Figure 4.2: Pictorial explanation of convolution; Assume we want to convolve two (discrete) functions  $f(x)$  and  $g(x)$  to obtain  $h(x) = f(x) * g(x)$ . To find the value of  $h(x_1)$  we have to shift the function  $g(t)$  in such a way that the origin of the dummy variable  $t$  is on the position of  $x_1$  then we have to multiply all the values  $f(x_1 - 3) \dots f(x_1 + 3)$  by  $g(-3) \dots g(3)$  respectively and then add them up.

separated in an even and odd part. How functions transform is shown in the diagram below.

$$\begin{aligned}
 f(x) &= f_o(x) + f_e(x) = \Re f_o(x) + j\Im f_o(x) + \Re f_e(x) + j\Im f_e(x) \\
 \downarrow \quad \downarrow & \quad \quad \quad \swarrow \quad \searrow \quad \downarrow \quad \downarrow \\
 F(s) &= F_o(s) + F_e(s) = \Re F_o(s) + j\Im F_o(s) + \Re F_e(s) + j\Im F_e(s)
 \end{aligned}
 \tag{4.6}$$

**SHIFT THEOREM** If the function  $f(x)$  is shifted horizontally in such a way that the new function becomes  $f(x + a)$ , then the Fourier transform of  $f(x+a)$  is the Fourier transform of the original function  $f(x)$  phase shifted by  $e^{2j\pi as}$ .

If  $f(x) \Rightarrow F(s)$  then  $f(x + a) \Rightarrow e^{2j\pi as} F(s)$

**CONVOLUTION THEOREM** The convolution of two functions  $f(x)$  and  $g(x)$  is another function  $h(x)$  defined as

$$h(x) \equiv \int_{-\infty}^{\infty} f(u)g(x - u)du
 \tag{4.7}$$

We denote the convolution of two functions by an asterisk (\*) e.g.  $h(x) = f(x) * g(x)$ . We refer to Bracewell, chapter 3, for more information about the properties of a convolution.

Convolution is a process in which a function  $f(x)$  is smoothed by a function  $g(x)$ . In figure 4.2 we give a pictorial explanation of the convolution process. An example of a convolution is shown in figure 4.3.

In radio interferometric imaging we encounter convolution when looking at brightness distributions smoothed by antenna patterns, when dealing with electrical signals and filters (e.g. effect of bandwidth), when digitally sampling data in the backends of the receiving system and in many other cases

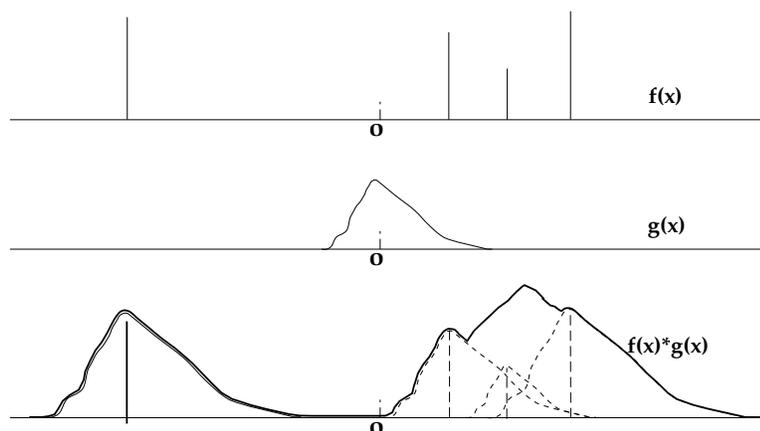


Figure 4.3: The convolution of a discrete function  $f(x)$  with a smooth function  $g(x)$ . The dashed lines are the convolutions of  $g(x)$  with the individual spikes.

The convolution theorem states that the Fourier transform of the convolution of two functions equals the product of the two Fourier transforms of these functions.

If  $f(x) \rightleftharpoons F(s)$  and  $g(x) \rightleftharpoons G(s)$  then  $f(x) * g(x) \rightleftharpoons F(s)G(s)$

**COMPLEX CONJUGATES** The complex conjugate of  $z = |z|e^{i\phi}$  equals  $z^* = |z|e^{-i\phi}$ . The Fourier transform of the complex conjugate of a function  $f(x)$  is  $F^*(-s)$ , that is, the ‘mirror image’ of the conjugate of the transform.

$$f^*(x) \rightleftharpoons F^*(-s) \tag{4.8}$$

**CORRELATION** The correlation of two functions  $f(x)$  and  $g(x)$  is another function  $h(x)$  defined as

$$h(x) \equiv \int_{-\infty}^{\infty} f^*(u-x)g(u)du = \int_{-\infty}^{\infty} f^*(u)g(u+x)du \tag{4.9}$$

where  $f^*(x)$  is the complex conjugate of  $f(x)$ . We denote the correlation of two functions by a small circle ( $\circ$ ) e.g.  $h(x) = f(x) \circ g(x)$ . In contrast to the convolution where  $f * g = g * f$ , the correlation function is not commutative

$$f(x) \circ g(x) \neq g(x) \circ f(x)$$

It can be shown that

$$f(x) \circ g(x) \rightleftharpoons F^*(s)G(s) \tag{4.10}$$

**SAMPLING THEOREM** If we sample a function we evaluate the function at discrete points. In fact we multiply the function by a series of delta functions separated by intervals  $\Delta x$ .

Sampling is encountered at many places in synthesis imaging. Some examples:

- In the backend of the receiver the output signal is measured at discrete time steps.
- The  $u, v$  plane is not completely covered because of the discrete spacing of the antennas.
- A computer calculates the Fourier transform by approximating the Fourier integral by a sum (cf. equation 4.5). The function under consideration is evaluated at discrete intervals.

The (Wittaker- Shannon or Nyquist) sampling theorem states that if a function  $f(x)$  is bandwidth limited *i.e.* the Fourier transform of  $f(x)$  vanishes for values of  $s$  outside the interval  $[-W, W]$ , then the function  $f(x)$  can be recovered completely if  $f(x)$  is sampled at a rate  $\Delta x$  for which:

$$\Delta x \leq \frac{1}{2W} \tag{4.11}$$

If the sampling rate does not satisfy this criterion (the Nyquist rate) aliasing will occur. This is illustrated in figure 4.4

The sampling theorem is derived in Thompson *al.* (1986) section 4.11, Gonzalez and Wintz (1987) section 3.3.9 and in chapter 10 of Bracewell.

## 4.5 REFERENCES

Bracewell, R.N. (1978): “*The Fourier Transform and Its Applications (2nd edition)*” McGraw-Hill International Book Company. ISBN 0-07-007013-X.

Gonzales, C. and Wintz, P. (1987): “*Digital Image Processing*” Addison-Wesley Publishing Company 1987 ISBN 0-201-11026-1.

Thompson, A.R., Moran, J.M. and Swenson Jr., G.W. (1986): “*Interferometry and Synthesis in Radio Astronomy*”. John Wiley & Sons, New York.. ISBN 0-471-80614-5. (see also review in part II Chapter 9.1, book 3).

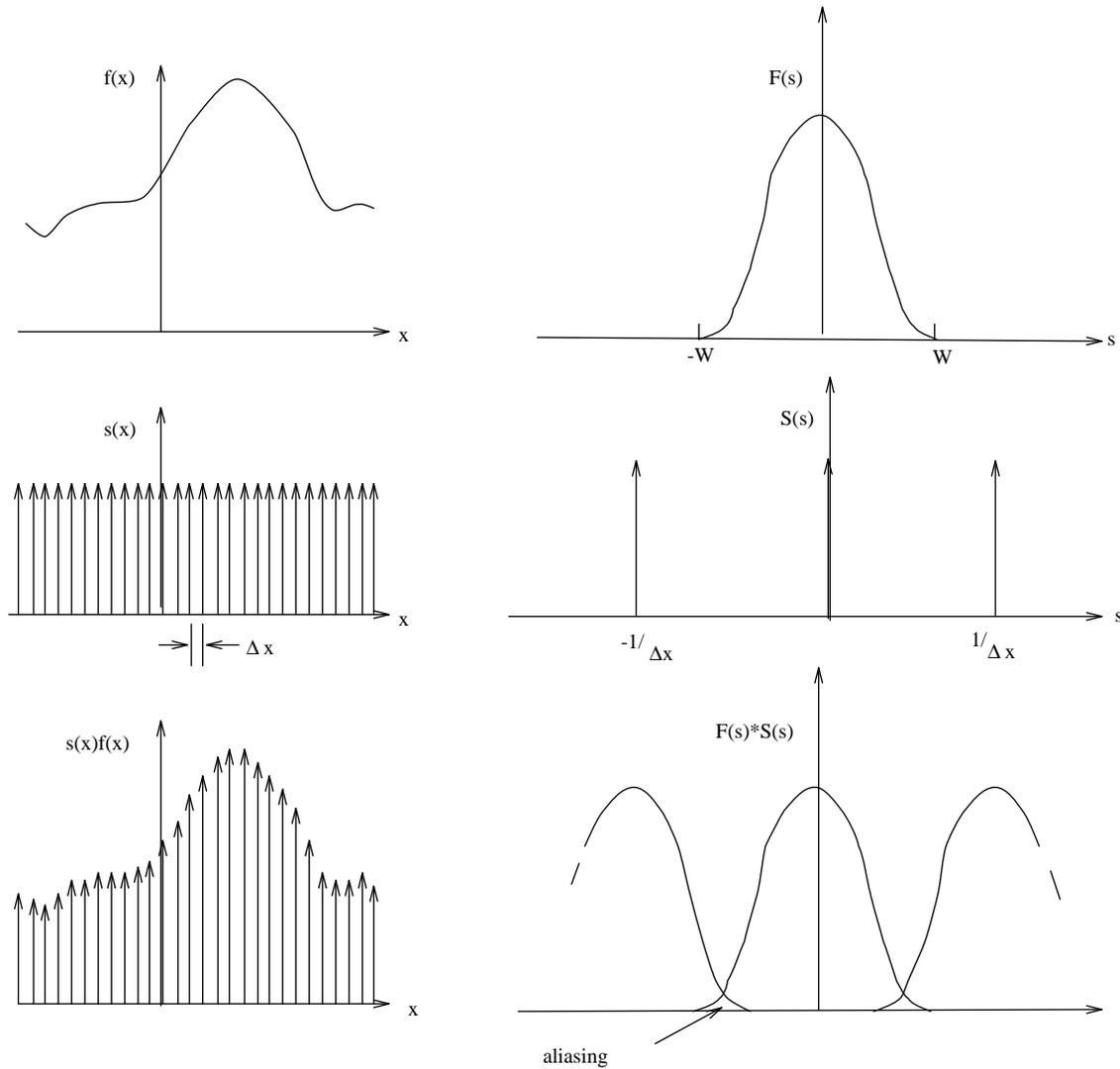


Figure 4.4: Sampling of a function  $f(x)$  is done by multiplying it with a sampling function  $s(x)$ , the process is shown in the left hand side of the diagram. In the Fourier domain the multiplication is equivalent to a convolution. The Fourier transform of the sampling function is convolved with the Fourier transform of  $f(x)$ , right-hand side of the diagram. The sampling theorem gives a limit on  $\Delta x$  for which there is no aliasing and in which case the function  $f(x)$  can be recovered from the data. Note that in this example aliasing will affect the high frequencies only because the function is nearly bandwidth-limited, this is not generally true. Synthesized images may contain aliased sources at any position.

## FAST FOURIER TRANSFORM EFFECTS

*Based on text from the first chapter of the old WSRT manual. That chapter was edited by A.G. Willis based on articles by J.A.Högbom, R.H. Harten and W.N. Brouw. The text has been re-edited and the figures have been redrawn by O.M. Kolkman*

In the preceding chapters we have shown how a brightness distribution can be obtained from a set of visibility measurements. We also reviewed the properties of the Fourier transform. In this chapter we discuss the artifacts introduced because the visibility must be sampled on a finite grid before a fourier transformation can be applied. We also discuss how these artifacts can be reduced and show what their appearance is in an astronomical image.

### 5.1 FAST FOURIER TRANSFORM EFFECTS AND A SUMMARY OF THE FUNCTIONAL RELATIONSHIPS INVOLVED IN APERTURE SYNTHESIS

In this chapter we will show how we compute Fourier transforms of our visibility function. Calculating the inverse of equation 2.1.(i.e equation 3.2) is accomplished by a discrete digital algorithm called the Fast Fourier Transform (FFT). The details of this procedure need not concern us here (refer to e.g. Bracewell (1986), chapter 18, or Gonzalez and Wintz (1987), section 3.4) but the FFT techniques require the visibility data  $\mathcal{V}(u, v)$  to be sampled at an evenly spaced rectangular grid of locations in the  $u, v$  plane. Instead we have measured  $\mathcal{V}(u, v)$  along a series of ellipses in the  $u, v$  plane. It is therefore necessary to interpolate the measured  $\mathcal{V}(u, v)$  onto a rectangular sampling grid by using a suitable convolving function, i.e. a function whose Fourier transform has negligible sidelobes outside the area of interest. Often a Gaussian or a prolate spheroidal function is used. We effectively distribute the values of  $\mathcal{V}(u, v)$  over the rectangular sampling grid with a weight that decreases as the distance of the rectangular sampling locations increases from the actual  $\mathcal{V}(u, v)$  measurement locations on the  $u, v$  ellipses.

We will now incorporate the effect of this extra convolution and the discrete Fourier transform into the following summary of data manipulation procedures which take place when we make a map of the sky by means of aperture synthesis.

Let us show what happens in a simple pictorial form. In this description we will use the notation of the previous subsections i.e. small letters represent functions in the  $u, v$  plane and capital letters represent their Fourier transforms in the sky plane ( $l, m$  coordinates). Also the dot symbol ( $\cdot$ ) represents multiplication, an asterisk ( $*$ ) represents convolution and a double arrow ( $\rightleftharpoons$ ) represents a Fourier transform. Thus  $a \cdot b \rightleftharpoons A * B$  represents the convolution theorem.

We define our functions as follows:

Functions in $u, v$ plane	Their Fourier transform
$b$ source complex visibility function (in figure figure 5.1 we only show the (constant) visibility amplitude)	$B$ sky brightness distribution (in figure 5.1 a point source)
$a$ average spatial frequency sensitivity function of a single antenna	$A$ average primary beam power pattern of an individual 25-m antenna
$c$ radial sampling function in $u, v$ plane	$C$ radial grating function
$t$ taper function in $u, v$ plane	$T$ taper function
$e$ convolution function to a rectangular $u, v$ grid	$E$ Fourier transform of the convolution function to the rectangular grid
$f$ $u, v$ plane rectangular grid sampling function	$F$ field repetition function

The functions are depicted in figure 5.1.

Note that  $\mathcal{V}(u, v)$  as defined in equation 2.1 equals  $b * a$  as defined here; the grating function  $g(u, v)$  as defined in section 5.2 equals  $c \cdot t$  and  $G(l, m)$  also defined in that section is essentially, but not quite, equal to  $C * T$ . The output map of the WSRT data is not merely a map of a region of the sky, but rather a map which has been convolved and multiplied by several sampling and convolution functions. We begin our observing process by pointing our antennas at a particular point in the sky. This region has a complex visibility function  $b(u, v)$ . We sample this distribution in two ways. First we limit our sky coverage by using highly directional antennas. Secondly we only measure the complex visibility at discrete increments (36 m, 72 m, ...), therefore, we actually observe a band of discrete spatial frequencies centered at each increment spacing. Over a 12 hour observation, these bands trace out ellipses in the  $u, v$  plane. Thus we are measuring the convolution of the complex visibility,  $b$ , and the antenna response,  $a$ , multiplied with a sampling increment function,  $c$ . This is the nature of the data as it is put on tape in Westerbork.

The data is now ready for off-line reduction. Initially it is calibrated and edited. (We will assume a perfect calibration for now, the effect of bad calibration is discussed in part IV chapter 3). After these corrections, the data is ready for the map making process. Three operations are performed on the data before it is Fourier transformed. The data is tapered by some taper,  $t$ , or grating function (see note above). This will tend to enhance the amplitude of certain spatial frequencies with respect to others. Then the data is convolved with a Gaussian or prolate spheroidal convolution function,  $e$ . This is necessary since, for the discrete Fourier transformation routine we need to sample the data in the  $u, v$  plane with an evenly spaced rectangular grid sampling function. After sampling we have the data in a form which can be Fourier transformed using the fast Fourier transform method. Clearly, the output of this transform will not just be a map of the sky brightness distribution, but the sky distribution multiplied or convolved with the Fourier transforms of the functions with which  $b(u, v)$  was convolved or multiplied respectively.

Let us examine what the data will look like at various stages of reduction. Assume we are observing a theoretical point source at the field center. The data, as it is collected at Westerbork, will consist of the source complex visibility convolved with the antenna visibility response multiplied by the sampling function along the baseline. This can be represented as

$$(b * a) \cdot c$$

Then after calibration a taper is applied to the visibility data; giving it a form

$$(((b * a) \cdot c) \cdot t$$

Then it is convolved to a rectangular grid in the  $u, v$  plane. Thus we have,

$$((((b * a) \cdot c) \cdot t) * e) \cdot f$$

The data is now Fourier transformed, using a fast Fourier transform method, yielding the output map

$$((((B \cdot A) * C) * T) \cdot E) * F$$

The effect of these different functions can be seen in figure 5.2.

The behavior of these functions in the sky plane can be described as follows.  $B$  is a delta function describing a theoretical point source.  $A$  is a general tapering of the map field by the single antenna response (*i.e.* the

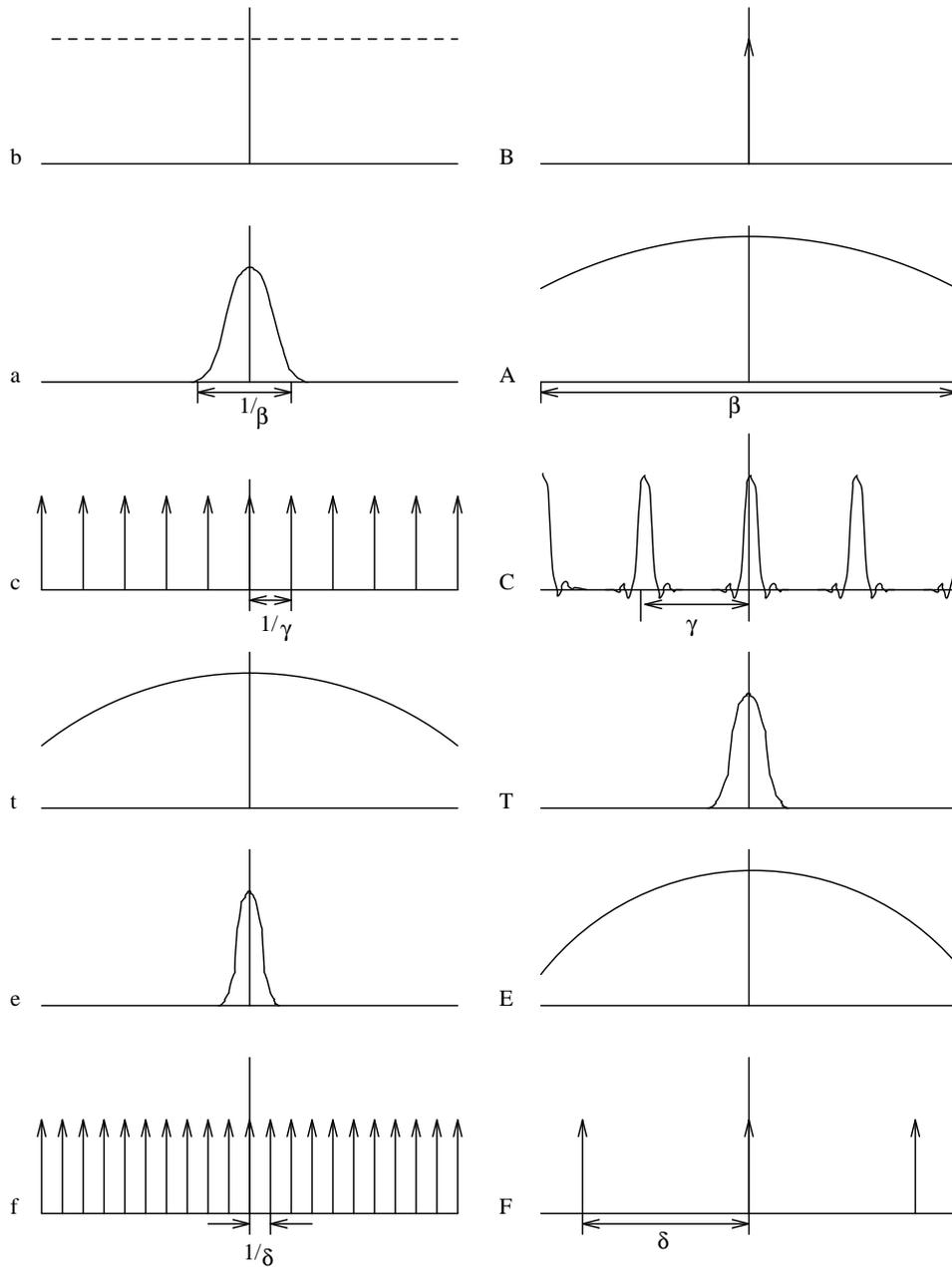


Figure 5.1: Functions involved in the map making process using the fast Fourier transformation, see text for explanation

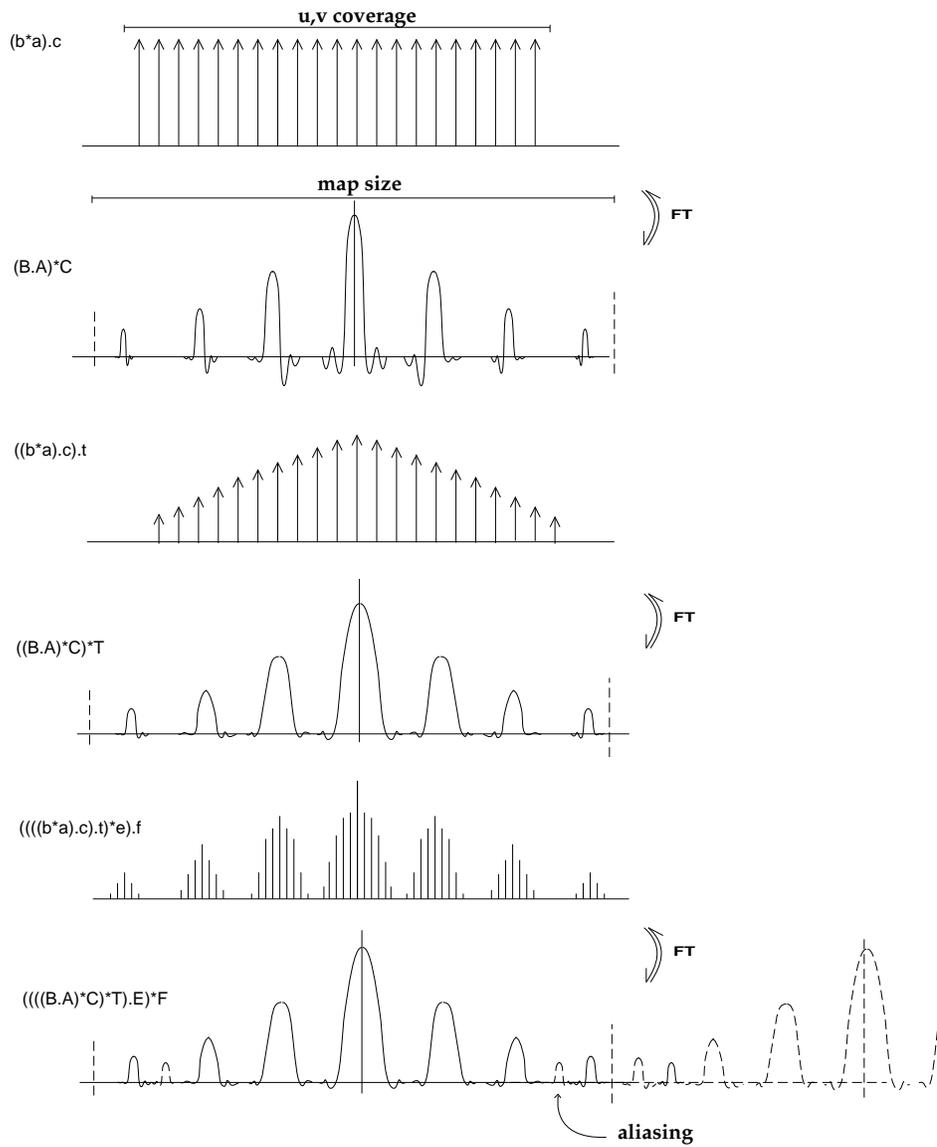


Figure 5.2: The effect of the different functions involved in map making using the fast Fourier transformation, see text for further explanation of the functions

primary beam) . Thus a point source would appear weaker , the further one moves it from the field center.  $C$  produces the grating rings.  $T$  smoothes the data in some manner to suppress the sidelobes of the synthesized beam.  $E$  also tends to taper the data and decreases the grating ring response as a function of distance from the source position. If  $E$  is a Gaussian of narrow width, (or order of half the field size), then the height of the grating rings will decrease quickly as the function of distance from the map center. Similar, if it is a broad Gaussian, then the amplitude of the grating rings will decrease slowly as a function of distance to the source position. This property of  $E$  is most important when considered in connection with the function  $F$ . The function  $F$  is the most bothersome one. It produces a repeating mosaic of the map field spaced on a rectangular grid whose points are separated one map width apart. We thus have not a single map but an infinite grid of the same maps. This might not seem important until one realizes that when we produce a map of the  $u, v$  data, we make a map of the entire sky. We only look at a small portion of this sky map in our  $512 \times 512$  or  $1024 \times 1024$  points output map. This is quite reasonable since the primary beam has limited our usable field to a small portion of the sky. The grating rings however are tapered mainly by the multiplication function  $E$  and may extend far beyond our usable field. Thus, any grating rings which might fall outside the field size of the map might appear in the adjacent map. This produces the effect of ‘reflection’. (They are not true reflections, but extensions from aliasing or an adjacent map.) To minimize this problem one should use a very narrow width Gaussian  $E$  function, which would cut down the intensity of the overlapping or reflected grating rings. The best solution would be to convolve the data to the rectangular grid with the aid of a convolving function  $e$  of the form  $\text{sinc}(x)$  (*i.e.*  $\frac{\sin x}{x}$ ), then the function  $E * F$  would be one in our primary map region and zero elsewhere. Unfortunately, a sinc function convolution is very expensive in computer time. Figure 5.3 demonstrates the effect of different convolving functions. Both the multiplication function  $E$  and the single antenna beam will attenuate the flux of a source displaced from the field center. To correct for this two things are done. First, the map is multiplied by the inverse of the multiplication function  $E$  in the sky plane. Thus the output map becomes,

$$((((B \cdot A) * C) * T) \cdot E) * F) \cdot (1/E) \quad (5.1)$$

This has the effect of correcting properly for the multiplication function within the map, except that the reflections do not have a proper correction and in general are too low. This is acceptable as long as you do not wish to remove them. Note that equation 5.1 is essentially the same as equation 3.2.

We can also correct for the primary beam attenuation by multiplying by  $1/A$ . We can retrieve the proper (convolved) sky brightness signal, but the noise everywhere in the map is also multiplied by  $1/A$ , which increases toward the edges of the map.

## 5.2 SYNTHESIZED BEAM AND GRADING

We have shown that during an observation the  $u, v$  plane is not covered completely. Before calculating the Fourier transform of equation 3.2 we introduce a grading function  $g(u, v)$  to weight the measurements. The grading  $g(u, v)$  is set to zero for all spacings  $(u, v)$  at which there are no measurements. The product of  $\mathcal{V}(u, v)g(u, v)$  is thus in contrast to the visibility  $\mathcal{V}(u, v)$  itself, known for all values  $(u, v)$ . Replacing  $\mathcal{V}(u, v)$  by this produces the integral of equation 3.2, we obtain from the convolution theorem in Fourier analysis

$$\{(B(l, m)A(l, m))\} * G(l, m) = \frac{1}{f(l, m, \delta_0)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{V}(u, v)g(u, v)e^{-2\pi j(u^l + v^m)} dudv \quad (5.2)$$

where  $G(l, m)$  is the Fourier transform of the grading  $g(u, v)$ . The convolution of  $G(l, m)$  with the expression within the curly brackets is equivalent to scanning the field with a telescope whose beam has the form  $G(l, m)$ . Thus this function, normalized to unity at maximum, will be called the synthesized beam.

The grading function,  $g$ , can be considered as a spatial frequency filter. The spatial frequencies in the brightness distribution are multiplied by this filter, so in the image domain the brightness distribution is convolved with the Fourier transform,  $G$ .

Let us for the moment consider a one dimensional slit. In optics the slit acts like a spatial frequency filter. The grading function for a slit will be a boxcar function. The brightness distribution will thus be convolved

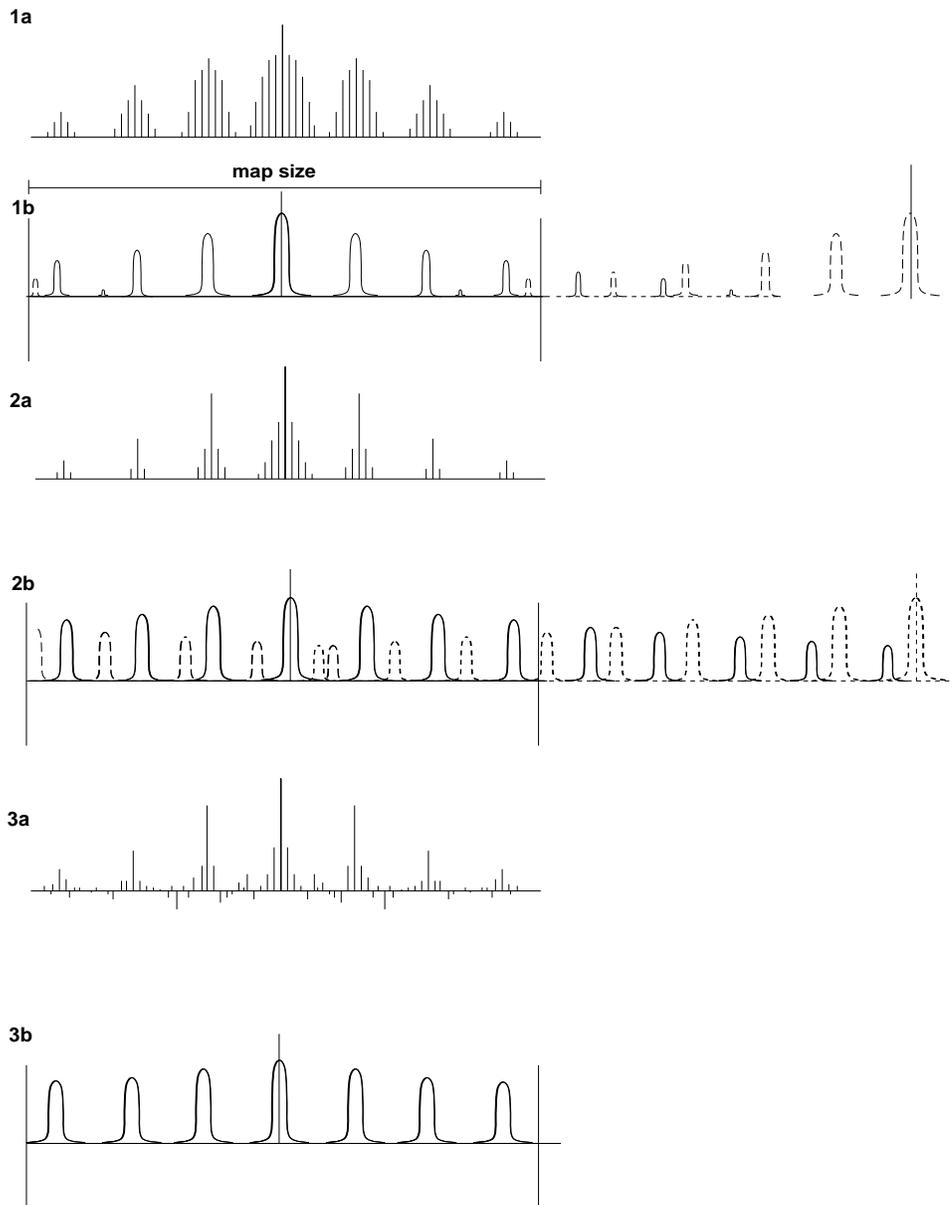


Figure 5.3: A well chosen weighting function (taper, grading) will reduce the power in the sidelobes. This reduces reflections or aliasing

with a sinc function (the Fourier Transform of a boxcar function). The intensity of a pattern projected by an uniformly illuminated slit is exactly this sinc function.

The synthesized beam is proportional to the Fourier transform of the grating  $g(u, v)$  whose form we may choose as we desire within the measured area of the  $u, v$  plane. A uniform grating ( $g(u, v)=1$ , where data was actually sampled) over this area will result in a beam with a -13% sidelobe. The sidelobes become smaller, at the expense of a somewhat wider synthesized beam, if the grating function is tapered towards the outer edge of the measured region so as to introduce a smoother transition from areas with data to areas without data.

The truncated Gaussian grating (tapered to 25% at the maximum baseline) used by default in Westerbork reduction programmes gives a -5% first sidelobe.

In angular measure, the shape of the synthesized beam is extended in declination by a factor  $1/\sin\delta$ . This is obvious also from the fact that the sampled  $u, v$  plane is only circular as viewed from the north celestial pole and becomes elliptical with the ratio  $1/\sin\delta$  when seen from other declinations.

The standard grating results in a synthesized beam whose width between half power points is  $0.8/D_\lambda$  radians in right ascension and  $0.8/(D_\lambda \sin\delta)$  radians in declination where  $D_\lambda$  is the maximum interferometer spacing in units wavelengths.

### 5.3 GRATING RESPONSES

During a 12h observation, the function  $\mathcal{V}(u, v)$  is only measured along a set of elliptical tracks in the  $u, v$  plane. The grating, as defined in the previous subsection, is equal to zero between these tracks and the smooth gradings discussed above give a simplified impression of the real situation. The finite number of measured tracks results in a synthesized beam pattern  $G(l, m)$ , given by the Fourier transform of the true grating, in which the central maximum is accompanied by a set of concentric grating ellipses. Expressed in radians these grating ellipses have semi-axes  $k/\Delta D_\lambda$  and  $k/(\Delta D_\lambda \sin\delta)$  radians in right ascension and declination directions, respectively where  $k$  is an integer and  $\Delta D_\lambda$  is the regular baseline increment in wavelengths. The amplitude of a grating is inversely proportional to the square root of its semi-minor axis. Thus, to minimize the disturbances caused by the grating ellipses, the true baseline increments  $\Delta D_\lambda$  should be small. A normal 12-h measurement taken with each group of two movable antennas separated by 72 m (half the spacing between the antennas in the fixed position array) yields an array in which the baseline is increased by regular 72 m increments.

This regularity then produces a set of elliptical grating responses whose semi-axes at a wavelength of 21 cm are multiples of 10 arc-minutes in right ascension and of  $10/\sin\delta$  arcmin in the declination direction. The dimensions of the ellipses, like those of the synthesized beam are proportional to the wavelength. In figure 5.4 a cross-section is shown of the synthesized pattern including the first two grating responses. Adding a second 12-h measurement with the movable antennas shifted by 36 m will give a combined array with a regular spacing of 36 m. This corresponds to grating ellipses with twice the previous size, *i.e.* all the odd numbered ellipses have been eliminated. after 2, 4, 8, etc. 12-h measurements with suitable positions of the movable antennas, the remaining grating ellipses will be 2,4, 8 etc times as distant as in the original set of ellipses.

### 5.4 ERRORS, SIDELOBES AND CONFUSION

Sidelobes of all kinds will degrade the synthesis map by giving rise to deflections on the map which are not at the position of the source which is their cause. Sidelobes appear as a consequence of the instrumental design, the choice of observational procedure and missing observations due to equipment malfunctions. The detailed shape of these can be calculated exactly and removed from the map by the CLEAN algorithm. Other map degenerations are caused by atmospheric phase fluctuations and by unavoidable, small departures of the instrument from its ideal calibrated performance like antenna dependent gain variations. These basically limit the dynamic range but it is often possible to estimate their rms amplitude distribution either from measurements or from a knowledge of the general stability of the atmosphere and the critical parts of the instrument. In principle the dynamic range can be improved by applying the SELFCAL (see e.g. the NewStar software descriptions) algorithm.

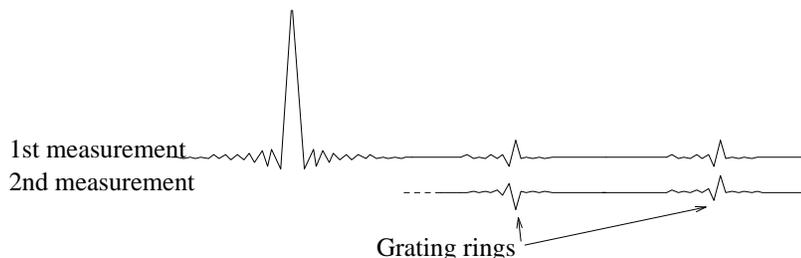


Figure 5.4: Schematically drawn cross-section of the synthesized beam pattern including the closest two grating rings for a single 12 h measurement when the two movable antennas are at 72 and 144 m respectively from the closest antenna of the fixed array. If instead the antennas are placed at 36 and 108 m respectively, the first and all other odd-numbered grating ellipses are reversed in amplitude as shown in the second cross-section. Combining the two sets of measurements, one can produce a map corresponding to a synthesized beam in which all the odd numbered rings have been eliminated. The remaining rings can be eliminated by making further measurements with the antennas in yet another settings of the movable telescopes.

Chapter 3 gives practical examples of maps distorted by calibration imperfections and other instrumental and atmospheric effects.

The set of grating ellipses discussed in the previous section is an example of the kind of sidelobe structure which can be calculated exactly. The problem of separating sources from sidelobes can become serious when the observed field contains sources which are larger in extent than the radius of the first grating responses (or when sources fall on the position of the grating rings). In such cases the grating disturbances must be eliminated, either by special data reduction procedures such as CLEAN (Högbom, 1974) or by adding further 12-h measurements.

No good quality measurements can be made at projected spacings which are smaller than the diameter of the individual antennas (25 m) because then one antenna would be blocking (or shadowing) part of the aperture of the other. Thus independent of the number of 12-h observations completed, there will always be a gap of missing spacings centered at the  $u, v$  origin. Its radius is a function of the smallest projected spacing actually used, but will usually be about 30m. The true grating of the synthesized aperture can therefore be written as:

$$g(u, v) = g_d(u, v) - g_o(u, v) \quad (5.3)$$

where  $g_d(u, v)$  is the desired complete set of ellipses in the  $u, v$  plane and  $g_o(u, v)$  represent those ellipses close to the origin that have not been measured. The synthesized beam  $G(l, m)$  is proportional to the Fourier transform of the grating and it follows that:

$$G(l, m) = G_d(l, m) - G_o(l, m) \quad (5.4)$$

*i.e.* the true synthesized beam equals the desired beam minus a beam which corresponds to measurements taken only at the missing small spacings. This latter is a broad low amplitude pattern. Thus, the global maximum of the true synthesized beam will be surrounded by an extended low level negative sidelobe structure. The integral over the entire synthesized beam pattern is equal to zero; this is a consequence of  $g(0, 0) = 0$  and the integral over a synthesis map must therefore also be zero. The negative sidelobe regions produce a depression

of the zero level which varies slowly over the map in a way which depends upon the detailed distribution and intensities of all sources in the field. This does not cause problems studying isolated small diameter sources, because the local zero level is sufficiently well determined by the surrounding empty parts of the map, but one has to be careful when calculating brightness temperature and flux densities of extended sources. These problems are avoided if the synthesis measurements are complemented by a survey of the same field obtained with a filled aperture telescope whose diameter is larger than the radius of the central gap in the  $u, v$  plane (see e.g. Rots, 1975). The extrapolation of large-scale structure in synthesis maps has also been described by Braun and Walterbos (1985)).

In some observations, notably at low declinations, it is unavoidable that shadowing of one dish by another occurs (see section 9). Shadowing is a problem because you can only make a proper correction for the field center. The effect of shadowing is always more on one side of the field than the other. A correction for the field center is a good approximation for a point source (e.g. a calibrator) but no use for an extended field. Many people simply delete shadowed data.

The dynamic range of the telescope is determined by the general sidelobe level caused by those effects — atmospheric fluctuations and instrumental instabilities— which cannot be calculated exactly. A weak source can only be determined if it is well above the random noise level on the map and the general sidelobe interference due to strong sources in the field. High dynamic range mapping can be achieved using special reduction techniques (e.g. Noordam and de Bruyn, 1982).

The term ‘confusion’ is usually employed in radio astronomy to refer to the fact that every observed field contains a large number of weak sources. These cause deflections that merge to a noise-like distribution over the map. For normal observations with the Westerbork telescope at 1415 MHz (or higher frequencies) this ‘confusion noise’ is below the sensitivity limit and has no influence on the interpretation of the synthesis maps. At 608 MHz, however, the greater flux density of most sources and the larger size of the synthesized beam combine to raise the confusion noise to a level greater than the sensitivity limit. The ‘confusion’ problem here is to decide how many deflections on a map can be interpreted as due to individual (point) sources rather than to a blend of many weaker sources. An often stated rule of thumb is to accept the largest deflections as individual sources but not to count more sources than what corresponds to an average of one source per 20-30 beam widths. The situation is even more complicated when the beam is accompanied by a prominent sidelobe pattern such as a set of grating rings. The statistical theory of the confusion and the influence of the detailed shape of the reception pattern have been discussed by Burns (1972). To reduce the problem of confusion more spacings should be measured to fill the  $u, v$ -plane better.

## 5.5 REFERENCES

- Bracewell, R.N. (1978): “*The Fourier Transform and Its Applications (2nd edition)*” McGraw-Hill International Book Company. ISBN 0-07-007013-X.
- Braun, R. and Walterbos, R.A.M. (1985): *Astron. & Astrophys.* , **143**, p. 307.
- Burns, W.R (1972): *Astron. & Astrophys.* , **19**, pp. 41.
- Gonzales, C. and Wintz, P. (1987): “*Digital Image Processing*” Addison-Wesley Publishing Company 1987 ISBN 0-201-11026-1.
- Högbom, J.A. (1974): *Astron. & Astrophys. Suppl.* , **15**, p. 417.
- Noordam, J.E. and de Bruyn, A.G. (1982): *Nature*, **299**, pp. 597-600.
- Rots, A.H. (1975): Ph.D. Thesis, University of Groningen.

# THE COORDINATE SYSTEM

*Ed. O.M. Kolkman based on text by W.N. Brouw (1972) and A.R. Thompson et al. (1986)*

After Fourier transforming a set of visibilities we obtain a brightness distribution. We now discuss the relation between the several coordinate systems used for Westerbork data.

## 6.1 $u, v$ AND $l, m$ COORDINATES

In this section we will introduce the  $(u, v)$  and  $(l, m)$  coordinate systems. The following discussion is only valid for east-west arrays. The equatorial coordinate system  $(x, y, z)$  is defined by the unit vectors:

- $\hat{e}_x$  which points in the direction  $\delta = 0^\circ, h = 0^h$ ,
- $\hat{e}_y$  which points in the direction  $\delta = 0^\circ, h = -6^h$  and
- $\hat{e}_z$  which points in the direction  $\delta = 90^\circ$ .

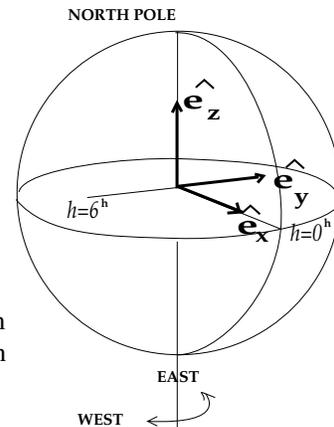
$\delta$  and  $h$  are the declination and hour angle respectively. In this coordinate system the baseline vector  $\mathbf{D}_\lambda$ , pointing from east to west with length  $D_\lambda$  can be written as

$$\begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = D_\lambda \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \quad (6.1)$$

If we point to a source in the direction  $(\alpha_0, \delta_0)$  (and associated hour angle  $h_0$ ) then the direction vector  $\mathbf{s}_0$  is given by

$$\begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix} = \begin{pmatrix} \cos \delta_0 \cos h_0 \\ -\cos \delta_0 \sin h_0 \\ \sin \delta_0 \end{pmatrix} \quad (6.2)$$

The geometrical delay  $\mathbf{D}_\lambda \cdot \mathbf{s}_0$  can thus be written as  $D_\lambda \cos \delta_0 \sin h_0$ .



The astrometric coordinate system is used for describing a tracking interferometer. The unit vectors of this orthogonal coordinate system are:

- $\hat{\mathbf{e}}_u$  which points from west to east as seen from the source,
- $\hat{\mathbf{e}}_v$  which points from south to north as seen from the source and
- $\hat{\mathbf{e}}_w$  which points from the source to the observer.

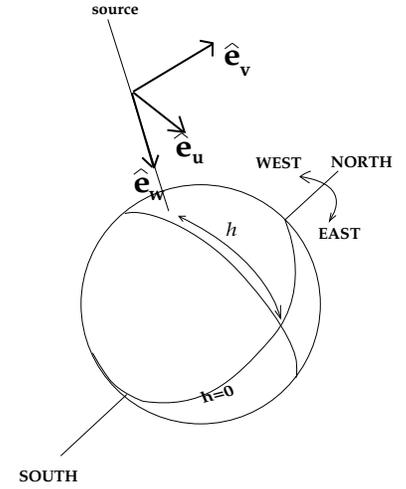
Let us first define  $D_u \equiv u$ ,  $D_v \equiv v$  and  $D_w = \mathbf{D}_\lambda \cdot \mathbf{s}_0$  (=the delay). Then the following transformation rules apply:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \sin h & \cos h & 0 \\ -\sin \delta \cos h & \sin \delta \sin h & \cos \delta \\ \cos \delta \cos h & -\cos \delta \sin h & \sin \delta \end{pmatrix} \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix}$$

From the above and eq. 6.1 we can immediately see that:

$$u = -D_\lambda \cos h_0 \tag{6.3}$$

$$v = -D_\lambda \sin h_0 \sin \delta_0 \tag{6.4}$$



When working with data from east-west interferometers it is common to define the coordinates  $l$  and  $m$  to describe the sky brightness distribution.

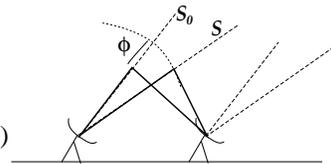
$$\sigma = \begin{pmatrix} l \\ m \end{pmatrix} \tag{6.5}$$

The origin of the  $(l, m)$  coordinate system is the field center  $\mathbf{s}_0 = (\alpha_0, \delta_0)$ . Below we will establish the relation between the  $(l, m)$  coordinates and the sky coordinates  $\mathbf{s}(\alpha, \delta)$ .

#### RELATION BETWEEN $(\alpha, \delta)$ AND $(l, m)$

Let us rewrite equation 2.1 in terms of  $(u, v)$  and  $(l, m)$  coordinates:

$$\begin{aligned} \mathcal{V}(\mathbf{D}_\lambda) &= \int_{4\pi} B'(\sigma) e^{-j2\pi\mathbf{D}_\lambda \cdot \sigma} d\Omega = \\ \mathcal{V}(u, v) &= \iint B'(l, m) e^{-j2\pi(ul+vm)} dl dm \end{aligned} \tag{6.6}$$



Now consider the phase term  $\phi = 2\pi(ul + vm)$  in the equation above. This phase term can be rewritten using the components of  $\mathbf{D}_\lambda$  (equation 6.3 and 6.4):

$$\phi = -2\pi D_\lambda (l \cos h_0 + m \sin \delta_0 \sin h_0) \tag{6.7}$$

Remember that for a delay tracking array this phase term is exactly the difference between the path length difference between two telescopes pointing at  $\mathbf{s}$  and the path length difference between two telescopes pointing at  $\mathbf{s}_0$ . *i.e.*  $\Delta\Phi = 2\pi(\mathbf{D}_\lambda \cdot \mathbf{s} - \mathbf{D}_\lambda \cdot \mathbf{s}_0)$ . The path length difference with respect to the direction  $(\alpha, \delta)$  equals  $2\pi D_\lambda \cos \delta \sin(h_0 - \alpha + \alpha_0)$  while the path length difference in the direction  $(\alpha_0, \delta_0)$  equals  $2\pi D_\lambda \cos \delta_0 \sin(h_0)$ , and hence we can also write the phase term  $\phi$  in the following way:

$$\phi = 2\pi D_\lambda (\cos \delta \sin(h_0 - \alpha + \alpha_0) - \cos \delta_0 \sin h_0) \tag{6.8}$$

From equation 6.7 and 6.8 we obtain the following relations between  $(\alpha, \delta)$  and  $(l, m)$ :

$$l = -\cos \delta \sin(\alpha - \alpha_0) \tag{6.9}$$

$$m = \cotan\delta_0 - \frac{\cos\delta \cos(\alpha - \alpha_0)}{\sin(\delta_0)} \quad (6.10)$$

The projection from  $(\alpha, \delta)$  to  $(l, m)$  is called the North Celestial Pole projection.

## 6.2 THE DIFFERENCE BETWEEN 3D AND E-W COORDINATES

It is important to note that the  $(u, v)$  coordinate system considered above is positioned in a plane with the pole as reference. This is an important simplification which can be made for E-W arrays. For 3D arrays one uses  $(u', v', w')$  coordinates. In the corresponding image coordinate system the point spread function of the beam is not constant while for our NCP coordinate system the PSF is constant over the field. For more details see Thompson et al.(1986) section 4.2 and 4.3<sup>1</sup>

## 6.3 REFERENCES

Brouw, W.N. (1971): Ph.D. Thesis, University of Leiden.

Thompson, A.R., Moran, J.M. and Swenson Jr., G.W. (1986): "*Interferometry and Synthesis in Radio Astronomy*". John Wiley & Sons, New York.. ISBN 0-471-80614-5. (see also review in part II Chapter 9.1, book 3).

---

<sup>1</sup>Note the notation in this document differs from the primed notation in this book

# TEMPERATURE, BRIGHTNESS AND SENSITIVITY

by O.M. Kolkman based on text from the books mentioned in section 9.1 and text from the old manual ed. by A.G. Willis

**Note:** We only discuss the theory here. For the calculation of sensitivity for the WSRT we refer to part III, chapter 2.

## 7.1 ANTENNA TEMPERATURE

In radio astronomy the concept of temperature is an important one. This has historical and practical reasons we will not get into here. The main assumption is that the characteristics of the signals involved are that of thermal noise from e.g. a resistor at a temperature  $T_r$ . The bandwidth limited noise power from such a resistor can be expressed in terms of its temperature using the Nyquist relation:

$$W_r = kT_r\Delta\nu \quad (7.1)$$

where  $W$  is the power,  $k$  is Boltzmann's constant,  $\Delta\nu$  is the bandwidth in which the power is emitted, and  $T_r$  is a temperature.

Using the Nyquist relation we can assign a temperature to the power measured by a radio antenna. Let us consider an antenna measuring a power per frequency interval  $W_A/\Delta\nu = w_A$ . ( $w_A$  is also called the spectral power.) Now the antenna temperature can be written as  $T_A = w_A/k$ . The antenna temperature is not only related to the physical temperature of the antenna but also to the temperature of the objects emitting radiation which heats the antenna. Details of these relations can be found in the standard text books discussed in section 9.1.

We define the effective aperture,  $A_e$ , of an antenna as a fraction,  $\eta_a$ , of the geometrical cross-sectional area *i.e.* the collecting area,  $A_T$ . The power per frequency interval received from a non-polarized source with brightness  $B(\theta, \phi)$  observed using an antenna with a effective aperture  $A_e$  and a (normalized) antenna pattern  $A_N(\theta, \phi)$  (see figure 7.1) can be written as:

$$w_a = \frac{A_e}{M} \int_0^{4\pi} B(\theta, \phi) A_N(\theta, \phi) d\Omega \quad (7.2)$$

where the factor  $M$  depends on the type of receiver and the dipole combination used. ( $M = 2$  in the case we are only sensitive for one dipole direction, thus half of the power received on the aperture is actually measured.)

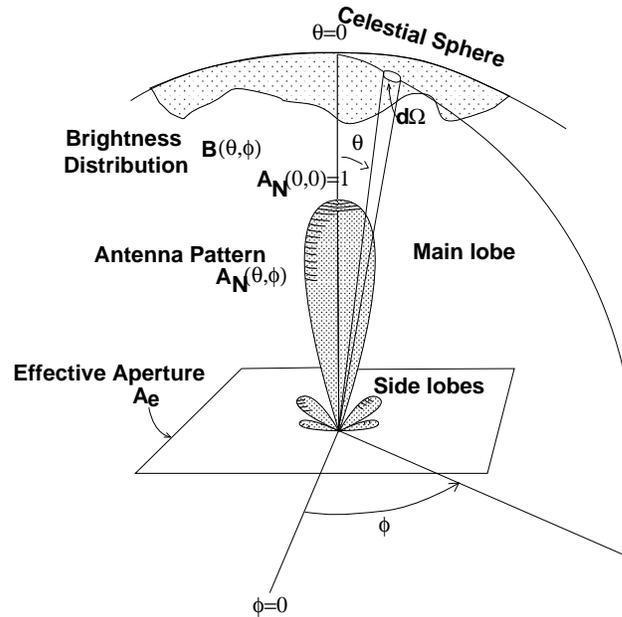


Figure 7.1: Diagram indicating the geometry of the antenna beam

From the equation above and the Nyquist relation we find that

$$kT_A = \frac{A_e}{M} \int_0^{4\pi} B(\theta, \phi) A_N(\theta, \phi) d\Omega \quad (7.3)$$

On the other hand we know that the flux density  $S_0$  of a source with brightness distribution  $B(\theta, \phi)$ , observed by the same antenna is given by:

$$S_0 = \int_0^{4\pi} B(\theta, \phi) A_N(\theta, \phi) d\Omega \quad (7.4)$$

$S_0$ , the flux density, has units of Watts  $m^{-2}Hz^{-1}$  or  $10^{26}$ Jansky.

From equation 7.4 and 7.3 we find that:

$$S_0 = M \frac{kT_A}{A_e} \equiv MS_m \quad (7.5)$$

where  $S_m \equiv kT_A/A_e$  is called the matched flux density.

Example: Consider a single 25 meter dish, measuring continuum radiation at 92 cm. At this wavelength the aperture efficiency  $\eta_a \approx 0.5$  so the effective aperture  $A_e \approx 12.5m$ . With a typical antenna temperature of  $T_A = 180K$  we find  $S_m \approx 2 \times 10^{-22}Watts m^{-2}Hz^{-1} = 2 \times 10^4Jansky$

## 7.2 BRIGHTNESS TEMPERATURE

The brightness temperature is related to the power emitted by a source. Assuming that the source radiates as a blackbody we can use the long wavelength approximation of the Planck law, the Rayleigh-Jeans Law:

$$B(\theta, \phi) = \frac{2kT_b(\theta, \phi)}{\lambda^2} \quad (7.6)$$

Note that the brightness temperature does not need to be equal to the physical temperature of the source. The relation between source temperature and brightness temperature depends on the emission mechanism. For further details we again refer to the standard text books discussed in section 9.1.

### 7.3 RELATIONS BETWEEN FLUX DENSITY, BRIGHTNESS AND TEMPERATURES

From equations 7.3 and 7.6 we find a relation between the antenna temperature,  $T_A$ , and the brightness temperature,  $T_b$ :

$$kT_A = \frac{A_e}{M} \int_0^{4\pi} \frac{2kT_b(\theta, \phi)}{\lambda^2} A_N(\theta, \phi) d\Omega \quad (7.7)$$

The relation between the matched flux density and the brightness temperature given by:

$$S_m = \frac{1}{M} \int_0^{4\pi} \frac{2kT_b(\theta, \phi)}{\lambda^2} A_N(\theta, \phi) d\Omega \quad (7.8)$$

In practice one assumes  $B(\theta, \phi)$  (and equivalently  $T_b(\theta, \phi)$ ) to be constant over an area larger than the area of the main lobe, in other words one assumes an extended source. In that case  $B(\theta, \phi)$  (or  $T_b(\theta, \phi)$ ) can be taken outside the integral and equations 7.7 and 7.8 reduce to

$$kT_a = \frac{A_e}{M} \frac{2kT_b}{\lambda^2} \int_{main\ lobe} A_N(\theta, \phi) d\Omega = \frac{A_e}{M} \frac{2kT_b}{\lambda^2} \Omega_A \quad (7.9)$$

and

$$S_m = \frac{1}{M} \frac{2kT_b}{\lambda^2} \int_{main\ lobe} A_N(\theta, \phi) d\Omega = \frac{1}{M} \frac{2kT_b}{\lambda^2} \Omega_A \quad (7.10)$$

respectively. The integral  $\int_{main\ lobe} A_N(\theta, \phi) d\Omega = \Omega_A$  is called the beam area. To make a proper estimate of the surface brightness or brightness temperature one ought to know  $\Omega_A$ . We will come back to this issue below.

In the case one is measuring a point source of brightness  $B$  or corresponding  $T_b$ , positioned at coordinates  $(\theta, \phi) = (0, 0)$  *i.e.* the beam center, the flux and temperature relations in equations 7.7 and 7.8 trivially reduce to

$$kT_a = \frac{A_e}{M} \frac{2kT_b}{\lambda^2}$$

and

$$S_m = \frac{1}{M} \frac{2kT_b}{\lambda^2}$$

respectively.

#### 7.3.1 ESTIMATES OF $\Omega_A$

The theoretical synthesized beam for a uniformly and completely filled aperture with Gaussian grading is

$$\Omega_A = 0.588362 \frac{\lambda}{\sin \delta D_{\lambda, max}} \frac{\lambda}{D_{\lambda, max}}$$

where  $D_{\lambda, max}$  is the length of the maximum Baseline in meters and  $\lambda$  is the wavelength in meters.

In practice it is most straightforward to actually measure  $\Omega_A$  by integrating the synthesized beam. If one works with CLEANed maps one must of course integrate over the restoring beam used by CLEAN to get  $\Omega_A$ . As this usually is a two-dimensional Gaussian beam one can evaluate  $\Omega_A$  analytically.

For a restoring beam:

$$A_N(\Omega) = A_N(l, m) = e^{-\left[\left(\frac{l}{b_l}\right)^2 + \left(\frac{m}{b_m}\right)^2\right] \ln 2} \quad (7.11)$$

where  $l, m$  are sky coordinates and  $b_l, b_m$  are the half power beamwidths in radians. Integrating over this beam yields:

$$\Omega_A = \int_{\Omega} A_N(\Omega) d\Omega = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left[\left(\frac{l}{b_l}\right)^2 + \left(\frac{m}{b_m}\right)^2\right] \ln 2} dl dm = \frac{\pi}{4 \ln 2} b_l b_m = 1.13 b_l b_m \quad (7.12)$$

### 7.3.2 CONVERSION FACTOR $T_b(\text{K})$ – $S(\text{mJy})$

Remember the relation between observed flux  $S$  and source brightness temperature,  $T_b$

$$S = \frac{2kT_b}{\lambda^2} \int_{\text{main lobe}} A_N(\theta, \phi) d\Omega = \frac{2kT_b}{\lambda^2} \Omega_A \quad (7.13)$$

Using the estimates for  $\Omega_A$  found above we can find the conversion factor between the flux density and brightness temperature

Using

$$\Omega_A = 0.588362 \frac{\lambda}{\sin \delta D_{\lambda, \max}} \frac{\lambda}{D_{\lambda, \max}}$$

where  $D_{\lambda, \max}$  is the maximum baseline sampled and  $M = 1$  (*i.e.* both dipoles used), we find

$$\frac{T_b(\text{K})}{S(\text{mJy})} = \frac{D_{\lambda, \max}^2 \sin \delta}{1.62 \times 10^6} \quad (7.14)$$

Note that the relation above does not depend on wavelength and is valid for line and continuum observations.

For an 2.75 km baseline we find  $T(\text{K}) = 4.68 \sin \delta S(\text{mJy})$ .

In the case of a two dimensional Gaussian (restoring) beam we find with  $M = 1$

$$S(\text{W}) = \frac{2k}{\lambda^2} \frac{\pi}{4 \ln 2} b_l b_m T_b(\text{K}) \quad (7.15)$$

Or rewritten with  $S$  in mJy, HPBW  $b_x$  and  $b_y$  in arcsec and substituting  $\lambda\nu = c = 2.997 \times 10^8 \text{m/s}$  and  $k = 1.380662 \times 10^{-23} \text{J/K}$

$$T_b(\text{K}) = \frac{1.2221261 \times 10^{22}}{\nu^2 b_x b_y} S(\text{mJy}) \quad (7.16)$$

And for the 21 cm wavelength region we find

$$\frac{T_b(\text{K})}{S(\text{mJy})} = \frac{605.7383}{b_x b_y} \left(\frac{\nu_0}{\nu}\right)^2 \quad (7.17)$$

where  $\nu_0 = 1420.405 \text{MHz}$ , the rest frequency of the 21 cm emission line.

## 7.4 SENSITIVITY

Note: only the theory is discussed here. For the calculation of sensitivity for the WSRT we refer to part III, chapter 2.

The sensitivity of a radio telescope is usually discussed in terms that suggest that all sources of error in the measurement are caused by random fluctuations with the well-defined statistical properties of thermal noise.

This assumption is valid only for ideal observing conditions — no interference, no scintillation and complete stable equipment. The sensitivity is a measure of the weakest source which, in the absence of confusing

sources, can be detected with confidence. It is often defined as the signal corresponding to the rms error detection,  $\sigma$  or as some multiple (usually between 2 and 5) of this number. If, for example, one wishes to determine whether a particular optical object is a radio source, then a positive detection of  $\approx 2\sigma$  at that position may be considered significant. The probability of a  $2\sigma$  positive detection occurring at any specified position is only about 1/44. If on the other hand one wants to discover previously unknown sources over a whole Westerbork field then this factor of 2 is not sufficient: the field of view is  $4 \times 10^4$  larger than the synthesized beam and one expects more than 800 positive detections of  $2\sigma$  or larger to occur there purely by chance even if the entire field is devoid of sources. The same level of confidence (1/44), in this case, can only be achieved by accepting detections which are  $4.6\sigma$  or larger. The sensitivity for this second program is therefore 2.3 times worse when the same map with the same rms error is used for both programs.

To calculate the sensitivity one needs to evaluate the noise fluctuations in the image how this is done is shown below.

Let us assign a system temperature  $T_s$  corresponding to the noise power of the receiver itself and all the other sources of unwanted noise.<sup>1</sup> The power of this system noise is much larger than the power in the cosmic signal. The r.m.s. fluctuations in the system noise power in terms of temperature,  $\Delta T$  integrated in a bandwidth  $\Delta\nu$  for a time  $t$  are proportional to  $(\Delta\nu t)^{-1/2}$  and can be expressed in units of system temperature,  $T_s$ , by

$$\Delta T = \frac{CT_s}{\sqrt{\Delta\nu t}} \quad (7.18)$$

where  $C$  is a factor that depends on the particular type of receiver ( $C = \sqrt{2}$  for the Westerbork correlation receiver.)

Using equation 7.5 we can find the r.m.s. error,  $\sigma$ , in our determination of the matched flux density.

$$\sigma(S_m) = \frac{k}{A_e} \frac{CT_s}{\sqrt{\Delta\nu t}} \quad (7.19)$$

In the case of an aperture synthesis telescope additional correction factors come in but the relation for the r.m.s. error in the matched flux density  $\sigma(S_m)$  is essentially the same as above. When the number of telescopes increases the noise in the system will decrease by  $1/\sqrt{N_I}$  and because a correlator may introduce some noise a degradation factor  $D$  comes in,  $\eta_g$  is a factor that arises from the fact that not all of the  $N_I$  telescopes have equal weight.

$$\sigma = \frac{D}{\eta_g \sqrt{N_I}} \frac{k}{A_e} \frac{CT_s}{\sqrt{\Delta\nu t}} \quad (7.20)$$

We now introduce the polarization summation factor,  $P_s$ . It is equal to one over the squareroot of the number of polarization channels  $N_P$ ,  $P_s = 1/\sqrt{N_P}$ .<sup>2</sup> Using this factor we can extract a general formula for the r.m.s. fluctuations,  $\Delta S$ , in the measured flux,  $S$ .

$$\Delta S = M P_s \sigma = \frac{M}{\sqrt{N_P}} \frac{D}{\eta_g \sqrt{N_I}} \frac{k}{2\eta_a A_T} \frac{T_s \sqrt{2}}{\sqrt{\Delta\nu t}} \quad (7.21)$$

where the effective aperture,  $A_e$  can be rewritten in terms of the geometrical cross-sectional area,  $A_T$ , of an individual telescope aperture and the antenna efficiency  $\eta_a$ :

$$A_e = 2\eta_a A_T$$

The other relevant terms in equation 7.20 are:

<sup>1</sup>In general the noise power from the receiver system is the largest. However, strong sources in the field can contain a lot of noise; especially at meter wavelengths (at which WSRT is going to observe in the future) the galaxy contains a lot of noise

<sup>2</sup>In Westerbork  $N_P = 1$  if we measure in XX or YY dipole setting,  $N_P = 2$  if we measure XX and YY dipole setting

$D$  = a degradation factor introduced by the digital correlation process used at Westerbork which raises the noise above that of a pure analogue correlation device.

$\eta_g$  = a grading efficiency factor arising from the fact that unequal weights are generally given to different interferometer pairs by the grading function applied to the data before Fourier transformation,  $\eta_g \approx 0.9$ .

$T_s$  = the system temperature. For an interferometer pair, the system temperature is given by

$$T_s = \sqrt{(T_F + T_A)(T_M + T_A)}$$

where  $T_F$  and  $T_M$  are the noise of the fixed telescope and the movable telescope and  $T_A$  is the antenna temperature due to the noises in the field.  $T_A$  is closely similar for all fixed and movable telescopes.

$N_I$  = the number of interferometers. For a  $N$ -element interferometer this number has a maximum of  $N(N - 1)/2$

$\Delta\nu$  = noise equivalent bandwidth of the observation. In the case of a continuum observation this value is approximately equal to the total bandwidth  $B$  of the observation. For line observations we are usually more interested in the noise per frequency channel. ( $\Delta\nu = B/N_F$  for a uniform and  $\Delta\nu = 2.67B/N_F$  for a Hanning taper respectively, where  $N_F$  is the number of frequency channels.)

Usually we want to know the total flux density,  $S$  and its associated error,  $\Delta S$ . They can be calculated from the matched flux density,  $S_m$ , and its error,  $\sigma$ , by substituting the instrumental parameter,  $M$ , which is now dependent on the relative position of the dipoles, and considering how the error changes if we add channels to obtain total flux.

Two examples:

**EXAMPLE 1** We measure only one polarization channel ( $N_P = 1$ ) and with parallel dipole pairs in the two antenna elements comprising an interferometer. The source flux density,  $S$ , then equals twice the matched flux density,  $S_M$  (so  $M = 2$ ).

The observed flux in one polarization channel (say XX)  $S_M = \frac{1}{2}S \pm \sigma$ . Normalization for channel flux gives a response  $S \pm 2\sigma$  Thus the r.m.s. error in the measured flux density is:

$$\Delta S = 2\sigma = \frac{D}{\eta_g \sqrt{N_I}} \frac{k}{\eta_a A T} \frac{T_s \sqrt{2}}{\sqrt{\Delta\nu t}}$$

where we have substituted  $M = 2$  and  $N_P = 1$  in equation 7.21.

**EXAMPLE 2** We measure in 2 channels and parallel dipoles ( $M = 2$ ). The total unpolarized source flux density equals  $S$ . Observed flux density in the (XX) channel =  $\frac{1}{2}S \pm \sigma_{XX}$ , in the (YY) channel the observed flux density equals  $\frac{1}{2}S \pm \sigma_{YY}$

Averaging these signals we obtain the measured flux density  $S'$  where

$$S' = \frac{1}{2} ((S \pm 2\sigma_{XX}) + (S \pm 2\sigma_{YY}))$$

or, when adding the noise quadratically ( $\sqrt{4\sigma_{XX}^2 + 4\sigma_{YY}^2} = 2\sqrt{2\sigma^2} = 2\sigma\sqrt{2}$ )

$$S' = S \pm \sigma\sqrt{2}, \text{ assuming } \sigma_{XX} = \sigma_{YY} = \sigma$$

# EFFECTS IN FOURIER TRANSFORMED SPECTRA

by A.G. Willis and J.D. Bregman

## 8.1 INTRODUCTION

At Westerbork, visibility spectra are made by complex Fourier transformation of a time cross-correlation function into the spectral frequency domain.<sup>1</sup> Here we discuss some instrumental effects which influence spectra obtained by this procedure. The discussion is basically oriented toward practical interpretation of results obtained with the digital Fourier transform procedure used at Westerbork and is not meant as a theoretical introduction to the subject of power spectrum analysis. For more detailed discussion of digital Fourier transform and spectral analysis techniques the reader may wish to consult the reference list at the end of this chapter. A short summary of general Fourier transform relations is given in chapter 4.

We first define even and odd functions.  $f_e(t)$  is an even function if  $f_e(-t) = f_e(t)$ . The Fourier transform of an even function is an even function and is real.  $f_o(t)$  is an odd function if  $f_o(-t) = -f_o(t)$ . The Fourier transform of an odd function is an odd function and it is imaginary. We note that an arbitrary function can always be decomposed into a sum of an even and an odd function.

We measure a real cross-correlation function in the time domain in Westerbork. Thus since the function can be split up in the sum of even and odd functions, the Fourier Transform in the spectral domain will always be composed of a real part which is even and an imaginary part that is odd. This property affects especially the shape of the spectrum obtained when the phase across the frequency band is not zero (see below).

Before we sample the time cross-correlation function the radio signals are mixed down to video frequencies covering the range  $\nu = 0$  to  $\nu = B$  where  $B$  is the bandwidth of the observation. Increasing frequency channel number in an observed spectrum always corresponds to increasing video frequency *but* increasing video frequency *may not* correspond to increasing radio frequency as different mixing schemes are used for the various radio frequencies available at Westerbork.<sup>2</sup> The main part of this section gives the explicit relation between channel number and increasing or decreasing radio frequency.

The actual video bandpass is given in the schematic diagram in figure 8.1. At the low frequency side there is a steep filter with 3kHz FWHM. At the high frequency side the spectrum is smoothly filtered to gradually fall to zero at frequency  $B$ . The last  $\sim 7\%$  of the band has a normalized response of less than 0.5.

<sup>1</sup>On page 2-6 we mentioned the relevant transformation in equation 2.11

<sup>2</sup>For the DCB you can get the lowest frequency at the center

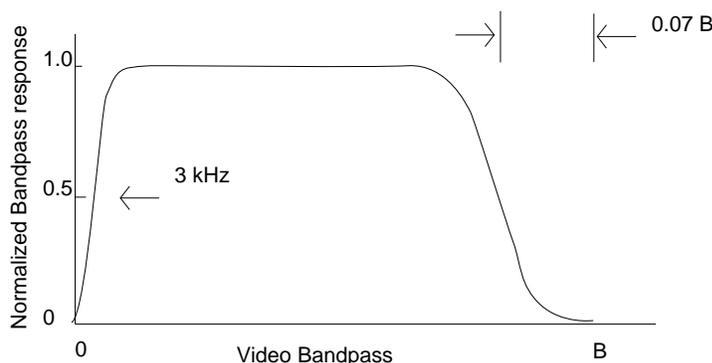


Figure 8.1: The bandpass

The video frequency is zero at all frequencies greater than B. The sampling theorem of Fourier analysis states that since  $f(\nu)$  is zero for all frequencies greater than B, then its Fourier transform, the cross-correlation function  $h(t)$  can be uniquely determined by discretely sampling at intervals  $\Delta t = \frac{1}{2B}$ . (Discrete sampling is obviously necessary for the performance of a digital Fourier Transform.)

At Westerbork the maximum bandwidth is 10 MHz, so the maximum sampling rate needed is  $\Delta t = 50$  nanoseconds. If we sample the time cross-correlation function at a rate slower than  $1/2B$  then we will introduce significant aliasing into the video spectrum that we calculate from the sampled cross-correlation function. We can see this effect from the schematic diagram (figure 8.2)

Referring to figure 8.2 we note that sampling the cross-correlation function at time intervals separated by  $\Delta\tau$  can be considered the same as multiplying the continuous correlation function by a sequence of  $\delta$  functions (or III-function, pronounce as Shah-function) separated by  $\Delta\tau$ .

We can Fourier transform the continuous function to yield the continuous frequency spectrum. Note that formally we must calculate values in the negative frequency domain although these are not "physically" meaningful. The Fourier transform of the time domain III-function is a frequency domain III-function whose  $\delta$ -functions are separated by  $\frac{1}{\Delta\tau} = 2B$ . Using the property that the Fourier transform of a function equal to two functions multiplied together (i.e. the sampled cross-correlation function) is equal to the convolution of the fourier transform of the two individual functions, we see that the Fourier transform of the sampled cross-correlation function is equal not to a single frequency, but a multiplicity of frequency spectra separated by  $2B$  in the frequency domain. Since the video frequency and its negative frequency image have a length of just  $2B$  we can see that we have avoided overlap of the positive frequency spectrum and its adjacent negative counterpart. If however we made the sampling time interval,  $\Delta\tau$ , longer we see that the III-function in frequency space would have its  $\delta$  functions closer together and thus the adjacent spectra forming the Fourier transform of the sampled cross-correlation function would begin to overlap and distort each other. This is the meaning of aliasing. (figure 8.3)

To reduce the aliasing effects, the Westerbork bandpass response is filtered so as to drop smoothly to zero at the maximum frequency B (see figure 8.1). We refer the reader to the book by Brigham (1974) for further details about aliasing.

Our next difficulty arises from the fact that the Fourier transform relation states that the frequency spectrum is related to the time cross-correlation by

$$f(\nu) = \int_{-\infty}^{\infty} h(t)e^{j2\pi\nu t} dt \tag{8.1}$$

i.e. we really need to sample  $h(t)$  out to time lags of  $\pm\infty$  before we can properly determine  $f(\nu)$ . Unfortunately, we are impatient scientists who do not wish to wait until the end of the universe before we determine the form of the spectrum  $f(\nu)$ . Thus in practice we only sample the cross-correlation function over some time range

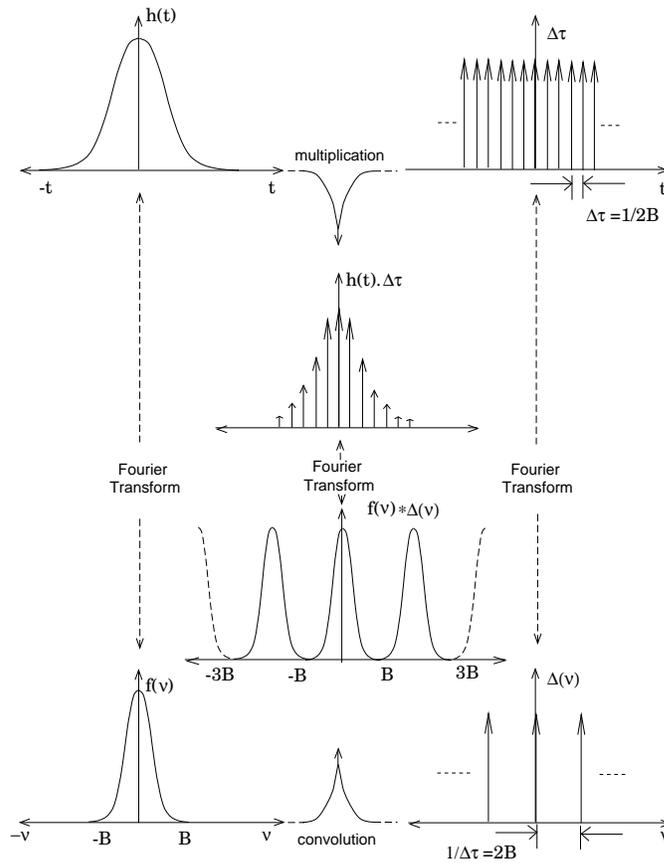


Figure 8.2: Fourier transform of a time function sampled at the rate  $\Delta\tau = \frac{1}{2B}$

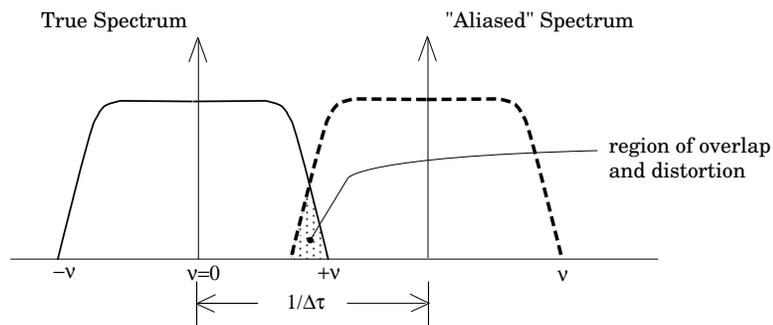


Figure 8.3: Aliasing

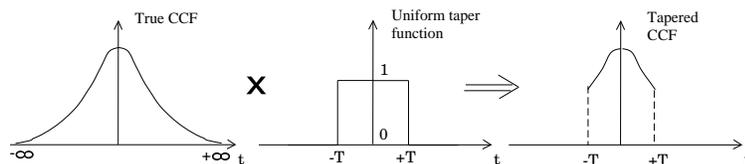


Figure 8.4: Tapering of the cross-correlation function (CCF) by a uniform taper

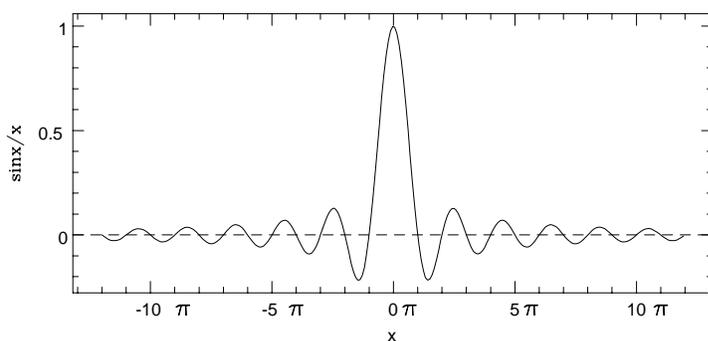


Figure 8.5: The  $\sin x/x$  function

$-T$  to  $+T$  before Fourier transforming the data. Our sampling of the cross-correlation function over only a limited time range can be considered the same as multiplying the cross-correlation out to infinite time lag by a *window function* (or taper) whose value is unity over the time range  $-T$  to  $T$  and whose value is zero beyond these bounds. This taper (which is the same as no taper) is what we mean by a uniform taper. The diagram in figure 8.4 illustrates this process in a schematic way.

Now remember that the Fourier transform of a function which is the product of two functions equals the convolution of the Fourier transform of the two individual functions. The Fourier transform of the uniform taper window is a  $\sin x/x$  function. Thus the spectrum we observe will equal the true spectrum (the Fourier transform of the proper cross-correlation function) convolved with a  $\sin x/x$  function.

We set the maximum value,  $T$ , out to which we sample the time cross-correlation function equal to  $N_F \Delta\tau$  where  $N_F$  is the number of frequency channels into which we wish to split up the frequency spectrum (which ranges from video frequencies 0 to  $B$ ) and  $\Delta\tau (= 1/2B)$  is the sampling rate of the cross-correlation function. Thus the total number of samples of the cross-correlation function that we need is  $2N_F$  since we actually sample the cross-correlation function from  $-T$  to  $+T$ , or over the time range  $2N_F \Delta\tau$ .

The Fourier transform of a window function of value unity over the range  $-N_F/2B$  to  $+N_F/2B$  and of value zero beyond these bounds is given (after normalization) by

$$f(\nu) = \frac{\sin \frac{2\pi N_F \nu}{2B}}{\frac{2\pi N_F \nu}{2B}} = \frac{\sin \frac{\pi \nu}{b}}{\frac{\pi \nu}{b}} \tag{8.2}$$

where  $b = B/N_F$  is the sampling interval in frequency space.

Looking at the  $\sin x/x$  function in figure 8.5 we can see why we made  $T$ , the maximum time out to which we measure the cross-correlation function, equal to  $N_F/2B$ . The resulting  $\sin x/x$  function given by equation 8.2 happens to have its null points, except for the central peak at frequency 0, lying exactly upon the multiples of

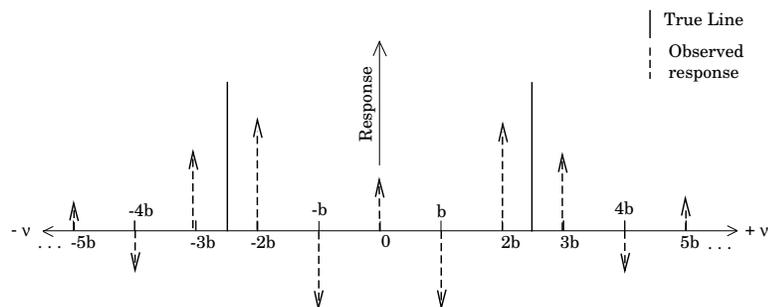


Figure 8.6: Observed response to a spectral line located between multiples of  $nb$

$b$ , where  $b$  is the sampling interval in the frequency spectrum.

Because the observed spectrum will be the true spectrum convolved by the function given in equation 8.2, the observed spectral response at frequency  $\nu_0$  to a true monochromatic frequency source of strength unity situated at frequency  $\nu_s$  is given by:

$$S(\nu_0) = \frac{\sin \frac{\pi(\nu_0 - \nu_s)}{b}}{\frac{\pi(\nu_0 - \nu_s)}{b}} \tag{8.3}$$

Which is just equation 8.2 except for a shift of the origin to frequency  $\nu_s$ . We then see that if the monochromatic source happens to lie at frequency  $\nu_s$  which is a precise multiple of  $b$ , the frequency sampling interval, then the observed spectrum will show a line only at frequency  $\nu_s$  and nothing elsewhere, since all the other sampling locations lie on nulls of the  $\sin x/x$  function.

We can extend this concept to show that if we have any number of monochromatic spectral lines, each of which is precisely located at a multiple of  $b$ , then the observed spectrum will still only show lines at precisely those locations and zeros at the other sampling locations.

Unfortunately most spectral lines are neither monochromatic nor do they occur at frequencies which are precise multiples of  $b$ . Because of the side-lobe characteristics of the  $\sin x/x$  frequency function we will measure non-zero frequency components at all discrete frequency sampling locations in the spectrum.

In the sample shown in figure 8.6 we show a true monochromatic line with phase  $0^\circ$  (remember we measure a complex frequency spectrum) at a frequency  $\nu_0$  which lies precisely between two sampling locations. The arrows show the magnitude and phase ( $0^\circ$  for positive directions,  $180^\circ$  for the negative directions) which we would measure for this line at different frequency sampling locations (remember that we must add in the response due to the "image" line in the negative frequency space). We could then use a fitting program to find the true position and strength of the line. However significant positive or negative amplitude responses are still measured at every frequency sampling location outside the main lobe of the  $\sin x/x$  function. This undesirable situation is called leakage; ways to reduce its effects are discussed below.

The full width half maximum (FWHM) of the  $\sin x/x$  function given by equation 8.2 equals  $1.2b$ . This gives the *effective frequency resolution* of the observation (*i.e.* two spectral lines whose frequencies are separated by less than  $1.2b$  will not be distinguishable as true separate lines). However, due to the significant near-in negative sidelobes, the noise equivalent bandwidth of a single frequency channel equals  $b$ .

We close this section by pointing out that the Fourier transformed spectral points situated at video frequency 0 can, obviously, have no phase information since its frequency is 0. The real amplitude of this spectral point is equal to the area under the cross-correlation function as will be apparent from equation 8.1. Since the data in this frequency channel (which is always frequency channel 0) is not astronomical useful, after the Fourier transform has been made the data in this channel are replaced by an *average* of the data in all the other frequency channels. This channel containing averaged data is then misleadingly referred to as the *continuum channel*

## 8.2 PHASE EFFECTS

Now consider what happens when we want to accurately calibrate a spectral line observation. To calibrate such an observation we separately observe a continuum point source whose spectral flux density,  $a$ , we assume to be perfectly constant across the total bandwidth,  $B$ , of the receiver. We can then use the observed response of the frequency sampling channels to calibrate the gain and phase corrections needed for each frequency channel.

At Westerbork we can apply an on-line phase zero correction to the data. However we will assume that the phase zero correction has not been perfect and that the visibility function of a continuum point source located precisely at the fringe stopping center still exhibits a residual phase  $\psi$  on a particular baseline.  $\psi$  is assumed to be constant across the bandwidth  $B$ .

To begin with, we will assume that the video bandpass has an idealized square shape; *i.e.* the response drops steeply to zero at frequency  $B$ . At frequencies less than  $B$  the response is assumed to be of constant amplitude. Thus in this idealized treatment we will initially ignore the actual amplitude tapering at the edges of the bands as shown in figure 8.1

Including the negative (and Hermitian) frequency responses, the spectrum of the continuum point source is then given by

$$f(\nu) = \begin{cases} ae^{j\psi} & \text{for } 0 < \nu \leq B \\ ae^{-j\psi} & \text{for } -B \geq \nu > 0 \\ a \cos \psi & \text{for } \nu = 0 \\ 0 & \text{for } |\nu| > B \end{cases} \quad (8.4)$$

The cross-correlation function  $F(t)$  of this function is given by

$$\begin{aligned} F(t) &= \int_{-\infty}^{\infty} f(\nu) e^{-j2\pi\nu t} d\nu \\ &= \int_{-B}^0 ae^{-j\psi} e^{-j2\pi\nu t} d\nu + \int_0^B ae^{j\psi} e^{-j2\pi\nu t} d\nu \end{aligned}$$

Remembering that an integral of the form

$$\begin{aligned} R(t) &= \int_x^y ae^{j\psi} e^{-j2\pi\nu t} d\nu \\ &= ae^{j\psi} \int_x^y (\cos 2\pi\nu t - j \sin 2\pi\nu t) d\nu \\ &= ae^{j\psi} \left\{ \frac{\sin 2\pi\nu t}{2\pi t} \Big|_x^y + j \frac{\cos 2\pi\nu t}{2\pi t} \Big|_x^y \right\} \end{aligned}$$

We can integrate the expression for  $F(t)$  to give

$$F(t) = 2aB \left[ \cos \psi \frac{\sin 2\pi Bt}{2\pi Bt} + \sin \psi \left( \frac{1 - \cos 2\pi Bt}{2\pi Bt} \right) \right] \quad (8.5)$$

We can see that  $F(t)$  contains a term  $\frac{\sin 2\pi Bt}{2\pi Bt}$  associated with the cosine of the phase and a term  $\frac{1 - \cos 2\pi Bt}{2\pi Bt}$  associated with the sine of the phase which is an odd function.

We now consider what would happen if the phase were exactly zero ( $\psi = 0$ ), *i.e.* we have a complete real and even signal of amplitude  $a$  in frequency space. Then the cross-correlation function would just be proportional to our friend the  $\sin x/x$  function which here has the form  $\frac{\sin 2\pi Bt}{2\pi Bt}$ . Now remember that we sample the cross-correlation function,  $F(t)$ , at time increments of  $\Delta\tau = 1/2B$ . We then see that excepting time  $t = 0$ ,

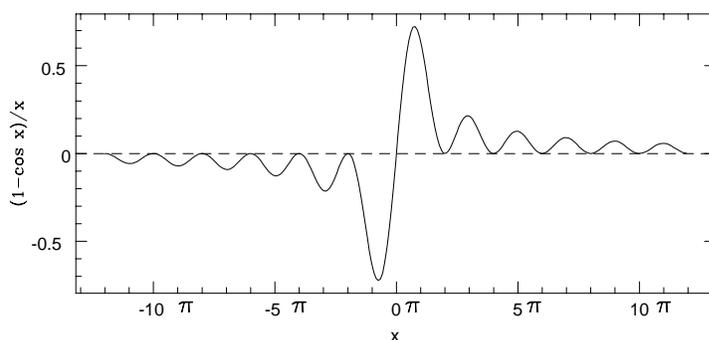


Figure 8.7: The  $\frac{1-\cos x}{x}$  function

we sample this cross-correlation function at precisely its null points. We only have measured response at  $2aB$  at time  $t = 0$  and zero elsewhere. This is just a  $\delta$ , or impulse, function.

We remark that in the case of discrete sampling, it is now necessary to multiply the impulse response by  $\frac{1}{2B}$  to obtain an impulse signal,  $a$ , which correctly represents the true area of the continuous  $\sin x/x$  function in the time domain. In the case of the continuous transform this is not necessary.

Since the measured cross-correlation function is zero except at  $t = 0$ , we can see that it is unimportant in this case that we stop measuring  $F(t)$  at a maximum time  $T = \frac{N_F}{2B}$  because the concept of a window (taper) function has become meaningless.

The fourier transform of the  $\delta$  of area  $a$  at  $t = 0$  is a function of constant amplitude  $a$  out to  $\pm\infty$  in frequency space, so we have effectively recovered the function given by equation 8.4 which just has constant amplitude  $a$  over the range  $-B$  to  $B$  when  $\psi = 0$

The effects on the sine component of the signal,  $-a \sin \psi$  for  $\nu$  less than 0 and  $a \sin \psi$  for  $\nu$  larger than zero, which is an odd function, are more difficult to calculate. When  $\sin \psi$  is non-zero, the time cross-correlation function contains a contribution from the term  $\frac{1-\cos 2\pi Bt}{2\pi Bt}$  which is also an odd function. Since the time sampling increment is  $\Delta\tau = \frac{1}{2B}$  we only measure  $\frac{1-\cos 2\pi Bt}{2\pi Bt}$  to have values of zero at every other sample point (see figure 8.7). At the sample points in between we measure  $\frac{1-\cos 2\pi Bt}{2\pi Bt}$  near or at absolute maximum. Thus it is clear that in this case when we measure  $\frac{1-\cos 2\pi Bt}{2\pi Bt}$  only out to limits of  $\pm \frac{N_F}{2B}$  we really have implicitly multiplied the function as it stretches out to  $\pm\infty$  by the uniform taper window function with value unity between the time limits  $\pm \frac{N_F}{2B}$  and value zero beyond these bounds.

Thus, when we transform the function  $\frac{1-\cos 2\pi Bt}{2\pi Bt}$  as measured between the limits  $-\frac{M_F}{2B}$  to  $\frac{N_F}{2B}$  back to frequency space we obtain our sine component of the  $2B$  frequency signal but it is now convolved with the transform of the window(taper) function. This transform was given by equation 8.2.

It is clear from equation 8.4 that the sine component of the signal has a discontinuous jump from  $-a \sin \psi$  to  $+a \sin \psi$  at video frequency 0. Thus our observed spectrum of the sine component of the signal will be the convolution of this sharp rectangular edge with the  $\sin x/x$  function given by equation 8.2.

To derive the resulting observed spectrum we shall neglect the fact that the value of  $|a \sin \psi|$  also jumps sharply to zero at  $|\nu| = B$  and assume that  $|a \sin \psi|$  is constant out to values of  $\nu = \pm\infty$ .

We justify this assumption by pointing out that the upper  $\approx 10\%$  of the actual video bandpass used at Westerbork is tapered to smoothly fall off to zero response at frequency  $B$ . Thus any effects in the observed spectrum due to convolution of the  $\sin x/x$  function with this smoothly tapered edge are negligible in comparison with the effects due to the convolution with the sharp edge assumed at frequency 0.

Because of this convolution, the observed sine component of the signal will exhibit a ripple pattern which is most dominant at frequencies close to the 0 frequency and which dies away as we move toward higher video frequencies. For the continuous Fourier transform, the observed sine signal is given as a function of increased

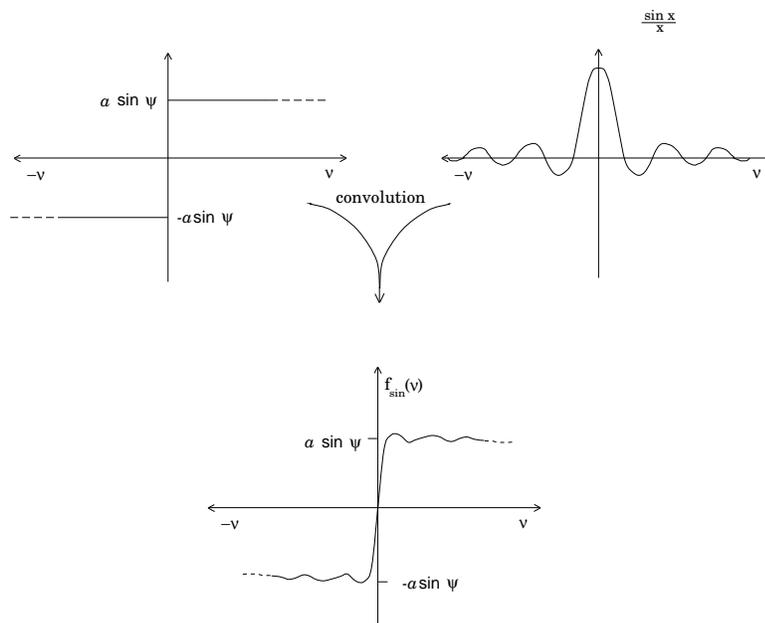


Figure 8.8: The result of the convolution of a sharp rectangular edge and a sinc function

observed frequency  $\nu_0$ , where  $\nu_0 > 0$ , (since we finally are interested only in the positive frequencies) by

$$\begin{aligned}
 f_{sine}(\nu_0) &= a \sin \psi \int_0^{\nu_0} \frac{\sin \frac{\pi \nu}{b}}{\frac{\pi \nu}{b}} d\nu \\
 &= a \sin \psi \frac{2}{\pi} \int_0^z \frac{\sin x}{x} dx
 \end{aligned}$$

if we change the variable of integration from  $\nu$  to  $x$ , and  $z = \frac{\pi \nu_0}{b}$ . We do the discrete Fourier transform, we sample the frequencies at discrete intervals of  $nb$  where  $n = 1, 2, 3, \dots$  and thus  $x = 0, \pi, 2\pi, \dots$ . Values of the function  $\int_0^z \frac{\sin x}{x} dx = Si(z)$  are tabulated by Abramowitz and Stengun (1965) in their chapter 5. The shape of the amplitude ripple is shown in detail in figures 8.8 and 8.9. We see that we will measure an  $f_{sin}(\nu_0)$  which can deviate from the value  $a \sin \psi$  by up to 18% (at the sample frequency point b).

Since the observed sine function has an amplitude ripple so will the observed phase as the observed phase is given by  $\arctan \left( \frac{\text{observed sin function}}{\text{observed cos function}} \right)$ . The measured amplitude clearly exhibits a ripple since it just equals  $\sqrt{\sin^2 + \cos^2}$ .

The magnitude of the observed ripples in amplitude and phase is clearly dependent on the phase  $\psi$  across the bandpass since only the term proportional to  $\sin \psi$  is affected by the convolution at a sharp edge.

In the above discussion we have neglected the actual filtering of the video bandpass on the low frequency side by the filter with a width of 3kHz at the half power level (figure 8.1). In situations where we only observe with a wide total bandwidth  $B$  and only a few frequency sampling points,  $b$  — the frequency sampling interval — is much greater than 3 kHz. In such case the frequency channel centered on video frequency 0 missed a fraction  $\frac{\approx 3}{b}$  of its power where  $b$  is expressed in kHz. So we effectively have an absorption feature centered on zero frequency much smaller than one channel width. Thus it is approximately a spectral point source which after deconvolution gives a response, observed at frequency  $\nu_0$ , of  $\frac{\approx 3}{b} S(\nu_0)$  which is given by equation 8.3 with

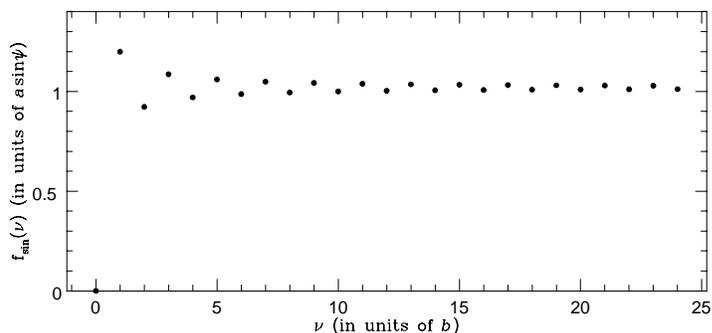


Figure 8.9: Values measured for  $f_{\sin}(\nu)$  at multiples of  $nb$

$\nu_s = 0$ . Since  $\nu_0$  is a precise multiple of  $b$  we sample the  $S(\nu_0)$  function at null points and no effect due to the absorption is seen (e.g. with  $B = 10\text{MHz}$  and  $N_F = 32$ ,  $b = 312.5\text{kHz} \gg 3\text{kHz}$ ). Thus for  $b \gg 3\text{kHz}$  the approximation used for the above calculation, that the video bandpass has uniform response completely across the band, is acceptable.

However, for small bandwidth observations with a large number of frequency channels, the 3kHz interval can be of the same order or greater than  $b$  (e.g. for galactic observations it is popular to select combinations of  $B$  and  $N_F$  such that  $b = 2.44\text{kHz}$ ). In such case the 3 kHz dip can extend over a few channels.

The amplitude of the ripple in the term proportional to  $\sin \psi$  will decrease because the transition from negative to positive signal is no longer sharply discontinuous at  $\nu = 0$  but smoother. However the cosine component will now drop down to zero at  $\nu = 0$ , *i.e.* we see an absorption wedge at frequency 0.

We will now see a ripple pattern in the term proportional to  $\cos \psi$  because of the convolution of this missing wedge with the  $\sin x/x$  function given by equation 8.2.

However the amplitude and shape of the ripple pattern in the  $\cos \psi$  function will be different to the amplitude and shape of the ripple in the  $\sin \psi$  function because in the first case the  $\sin x/x$  function is convolved with an absorption wedge while in the second case the convolution is with a ramp. Thus there will still be a ripple in amplitude and phase whose exact shape is dependent on the relative contributions from the convolved cosine and sine functions *i.e.* on the original phase  $\psi$  in the band.

Thus uniform (or no) taper observations should be approached with caution. The phase  $\psi_0$  of an observation can easily be different to the phase  $\psi_c$  of its calibrator because of additional extended structure, sources away from the field center, etc. Thus the observed amplitude and phase ripple of the frequency spectrum of an observation may be different to the observed amplitude and phase ripple of the calibrator. Consequently the observation will end up being improperly calibrated. However as we can see from figure 8.9, the ripple will have an amplitude of more than 2 percent in only the first twenty low numbered channels. Thus if we make a line observation with  $\geq 64$  frequency channels and the line we are interested in is located in only the central part of the band, then the uniform taper option is still a viable observing procedure if the line signal per frequency channel is considerably greater than 25% of the continuum signal.

### 8.3 TAPERING

The undesirable amplitude and phase ripples discussed above are caused by the measurement of the time cross-correlation function only out to some maximum time lags  $\pm T$ . When we only use uniform taper (= no taper, see above) we have sharp discontinuities in the cross-correlation function at time lags  $\pm T$

It are these discontinuities with their associated  $\sin x/x$  transforms which produce the undesirable amplitude and phase ripples in the frequency spectra.

We can reduce the effect of the discontinuities by giving the outer parts of the measured cross-correlation function reduced weight. (Remember that with uniform taper the weight can be considered unity everywhere

over the measured time lag and zero beyond the measured interval). At Westerbork we presently apply this reduced weight by means of a window function or taper called Hanning (after van Hann, an Austrian mathematician), which in the time domain has the form

$$f(t) = \begin{cases} \frac{1}{2} (1 + \cos(\frac{\pi t}{T})) & \text{for } |t| \leq T \\ 0 & \text{for } |t| \geq T \end{cases} \quad (8.6)$$

This function goes smoothly to zero at time  $T$ .

The Fourier transform of this window function in the frequency domain, or the observed spectral response at frequency  $\nu_0$  to a monochromatic point source situated at frequency  $\nu$ , is given by

$$F(\nu_0 - \nu) = \frac{1}{2}S(\nu_0 - \nu) + \frac{1}{4}S(\nu_0 - \nu + b) + \frac{1}{4}S(\nu_0 - \nu - b) \quad (8.7)$$

where  $S$  is the Fourier transformed uniform taper spectral response to a spectral point source (equation 8.1).  $\nu_0 - \nu$  is a multiple of the frequency sampling interval  $b$ .

$F(\nu_0 - \nu)$  has a full width half maximum of  $2b$ . Thus  $2b$  is the effective frequency resolution of a Hanning tapered observation. However the near-in negative sidelobes of the Hanning convolving function are only 2% of the peak response, in contrast to the 21% negative sidelobes of the uniform taper  $\sin x/x$  convolving function. Thus the Hanning convolving function is broader than the  $\sin x/x$  convolving function but has smaller negative sidelobes. Also we find that the amplitude and phase ripple in a Hanning tapered observation is very small (less than  $\sim 2\%$ ).

## 8.4 DELAY EFFECTS

The time cross-correlation of a visibility function is obtained by correlation of two antenna signals of an interferometer pair. Before the signals enter the cross-correlator they are fringe-stopped and delay corrected. The first correction eliminates the Doppler frequency difference between the two elements looking at a sky source from the rotating earth. The delay correction eliminates the path difference of the plane wave incident on the two elements of the interferometer.

In 1980<sup>3</sup> a digital delay system was implemented which corrects the video frequency signal. The delay will be accurate to within  $\sim 0.1ns$  in a  $10s$  integration period. Thus, with the proper fringe stopping for zero video frequency, there will be negligible phase slope oscillation over the video frequency band.

Delay errors can occur however and that is why when we observe HI at velocities close 0 we have to do a calibration observation below and above the frequency of interest. If the delays were perfect one calibrator would be OK, but if you shift frequency you usually find that the phase-zero offsets shift noticeably (typically  $5^\circ/MHz$ )

## 8.5 FAST FOURIER TRANSFORM EFFECTS

The fast Fourier transform (FFT) algorithm requires that the total number of sample points of the cross-correlation function and spectral frequency functions each be a discrete power of 2. This means that our sampling of these functions can not be precisely symmetric about the origins  $t = 0$  or  $\nu = 0$ . The time cross-correlation function is sampled at  $2N$  points separated by  $\Delta\tau = \frac{1}{2B}$ . The  $N$ th point is near  $t = 0$  and the so called odd channel near  $t = -N\delta\tau$ . Using the FFT the  $2N$  points of the cross-correlation function are

<sup>3</sup>Until 1980 the delay was corrected in steps of  $10m$  (33.3ns) at an intermediate frequency of  $30MHz$  in intervals of 2 minutes. This causes the cross-correlation function to be time shifted by a maximum of about  $20ns$ . According to the shift theorem for Fourier transforms (see section 4) this causes a phase slope over the band in the frequency domain. Through a change of the fringe stopping frequency, there is no phase change for the center of the band. After correlation every data sample is corrected for the average phase slope across the band during that integration period. Even with a good phase zero correction at the band center, this phase slope across the band will produce, for uniform tapered observations, amplitude and phase ripples in the frequency spectra similar to those discussed above for a constant phase offset. However, the difference between the phase at the edge of the band and the phase at the band center oscillates in amplitude (up to  $90^\circ$ : *i.e.* differences up to  $\pm 90^\circ$  for a  $10MHz$  bandwidth) as the delay offsets change. Thus the amplitude and phase ripple effects due to the delay offsets are averaged out when a frequency spectrum is made from a large number data samples taken with different delay settings.

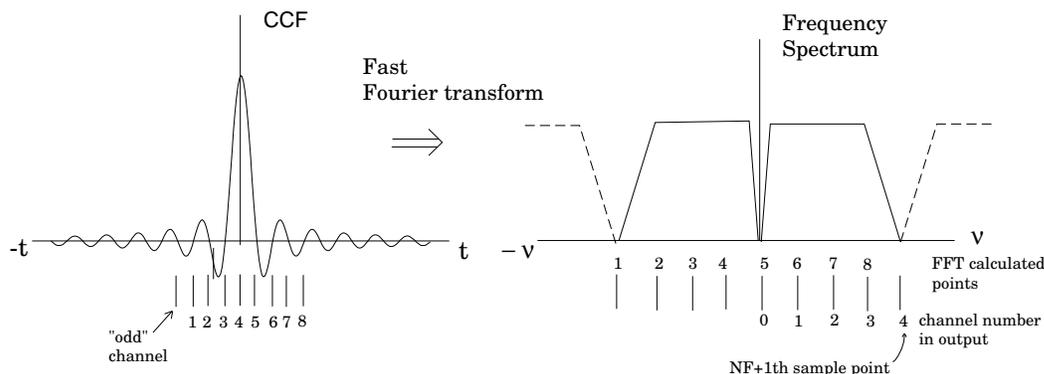


Figure 8.10: FFT relationships

transformed into complex frequency points of which  $N - 1$  are useful channels at video frequencies between 0 and  $B$ .

Remember that the data centered at video frequency has no phase and is not physically meaningful. The situation is sketched out in figure 8.10 for  $2N = 8$ , which would correspond to requesting a spectrum with  $N_F (= N)$ , the number of frequency points, being equal to 4.

The  $N_F$  output frequency channels are given by the numbers 0, 1, 2, 3, ...,  $N_F - 1$

Note that the  $N_F + 1$ th sample point would have a contribution from the true spectrum and the adjacent alias spectrum and thus does not contain meaningful data. This result does not affect frequency spectra measured with a uniform taper, but in a Hanning tapered observation, the  $N_F$ th point contains a 25% contribution from the  $N_F + 1$ th point (see equation 8.7 and 8.8) and must be discarded. (the data in this part of the spectrum has very low signal anyway because of the bandpass taper, so deleting the  $N_F$ th point we haven't lost useful data)

## 8.6 NOISE DISTRIBUTION IN TAPERED SPECTRA

As we can see from equation 8.7 use of the Hanning window function in the time domain corresponds to smoothing a frequency spectra Fourier transformed from a uniform tapered time correlation function with the operation

$$H_i = \frac{1}{4}U_{i-1} + \frac{1}{2}U_i + \frac{1}{4}U_{i+1} \tag{8.8}$$

where the  $U_i$  are the signals in the  $n + 1$  frequency channels (having numbers 0 through  $n$ ) of a frequency spectrum produced from uniform tapered CCF data and where the  $H_i$  are the signals in the  $n + 1$  channels of a frequency spectrum produced from Hanning tapered CCF data.

We now relate the noise in a Hanning tapered channel to the noise in a uniform tapered channel. Here we ignore the data in channels 0,1 and  $n$  of the Hanning tapered observation since they contain a contribution from the video frequency 0 or the aliased spectrum (see above).

Thus, if we define the noise in a single uniform tapered channel to be  $\sigma$  (assumed to be the same for all channels) the noise in the Hanning tapered channel is

$$\begin{aligned} \sigma_{H_i} &= \sqrt{\left(\frac{1}{4}\sigma_{i-1}\right)^2 + \left(\frac{1}{2}\sigma_i\right)^2 + \left(\frac{1}{4}\sigma_{i+1}\right)^2} \\ &= \sqrt{\left(\frac{1}{16} + \frac{1}{4} + \frac{1}{16}\right)\sigma^2} \\ &= \sqrt{\frac{6}{16}}\sigma = 0.61\sigma \end{aligned}$$

		Taper	
		Uniform	Hanning
channel spacing		$b = \frac{B}{N_F}$	$b$
FWHM (effective frequency resolution)		$1.2b$	$2b$
noise equivalent bandwidth per channel		$b$	$(8/3)b$
correlation between	first	0	2/3
	second	0	1/6
	third	0	0
Average noise per channel as a result of summing $n$ adjacent channels			
n=	one	$\sigma$	$\sigma(3/8)^{\frac{1}{2}}$
	two	$\sigma/\sqrt{(2)}$	$\sigma\frac{1}{2}\sqrt{(1+\frac{1}{4})}$
	three	$\sigma/\sqrt{(3)}$	$\sigma\frac{1}{3}\sqrt{(2+\frac{1}{4})}$
	four	$\sigma/\sqrt{(4)}$	$\sigma\frac{1}{4}\sqrt{(3+\frac{1}{4})}$
	⋮	⋮	⋮
	eight	$\sigma\sqrt{(8)}$	$\sigma\frac{1}{8}\sqrt{(7+\frac{1}{4})}$

Table 8.1: A summary of uniform and Hanning taper properties

Thus the noise equivalent bandwidth of a Hanning tapered channel =  $\frac{16}{6} = 2.67 \times$  that of an uniform tapered channel.

However the noise in adjacent Hanning channels is **strongly correlated**, so that summation and averaging of two adjacent Hanning tapered channels does not result in a  $\sqrt{2}$  decrease in the noise.

e.g. If we add together two adjacent Hanning channels, say  $i$  and  $i + 1$ , then using the relations between Hanning and uniform channels given by equation 8.8 we have

$$\begin{aligned}
 H_{AV}(\text{averaged signal}) &= \frac{1}{2}(H_i + H_{i+1}) \\
 &= \frac{1}{2} \left( \frac{1}{4}U_{i-1} + \frac{1}{2}U_i + \frac{1}{4}U_{i+1} + \frac{1}{4}U_i + \frac{1}{2}U_{i+1} + \frac{1}{4}U_{i+2} \right) \\
 &= \frac{1}{8}U_{i-1} + \frac{3}{8}U_i + \frac{3}{8}U_{i+1} + \frac{1}{8}U_{i+2}
 \end{aligned}$$

Assuming the noise,  $\sigma$ , to be the same in each of the uniform tapered channels, we find that

$$\begin{aligned}
 \sigma_{AV} &= \sqrt{\frac{1}{64}\sigma^2 + \frac{9}{64}\sigma^2 + \frac{9}{64}\sigma^2 + \frac{1}{64}\sigma^2} \\
 &= \frac{\sqrt{20}}{8}\sigma = 0.56\sigma
 \end{aligned}$$

Remembering that the noise in 1 Hanning channel was  $0.61\sigma$  we see that the addition of two adjacent channels has resulted in a noise decreased by only a factor  $\frac{0.56}{0.61} = 0.92$  instead of the factor  $\frac{1}{\sqrt{2}} = 0.71$  expected for two completely independent signals.

If we add together two Hanning channels separated by  $2b$ , e.g. numbers  $i$  and  $i + 2$ , then by the same procedure as outlined above we find that the decrease is 0.82, so there is still some correlation (As is to be expected since uniform taper channel  $i + 1$  contributes  $\frac{1}{4}$  signal to each of the Hanning channels  $i$  and  $i + 2$ .)

Properties of uniform and Hanning tapered spectral channels are summarized in table 8.1.

## 8.7 OTHER TAPERS

As shown above sidelobes in the frequency domain can be suppressed by tapering of the cross-correlation function in the time domain. This can be done by a number of tapering functions. These are discussed by Harris (1978). Three taper functions can be used at Westerbork. They are all applied before the CCF is Fourier transformed. We already discussed the Uniform (no taper) and Hanning taper above. The Hamming taper is similar to the Hanning taper but has a narrower frequency beam and lower side lobes. For the Hamming taper the individual frequency channels are also correlated in a similar way as for the Hanning taper (see previous section). The sidelobes of the Hamming taper do not decrease in amplitude as rapidly as those of the Hanning taper so the correlation involves more channels than only the adjacent, as is the case for the Hanning taper. The differences between the Hanning and Hamming taper are indicated in the table below:

	W(i)	Beamwidth	Highest side lobe	
		FWHM	in db	in %
Uniform		$1.21 B/N_F$	-6.5	22.4
Hanning	$0.5 + 0.5 \cos(2i\pi/N)$	$2.0B/N_F$	-16.0	2.5
Hamming	$0.54 + 0.46 \cos(2i\pi/N)$	$1.81B/N_F$	-21.5	0.7

where:

$W(i)$  is the discrete form of the function in the time domain

$$i = -\frac{N}{2}, -\frac{N}{2} + 1, \dots, -1, 0, 1, \dots, \frac{N}{2} + 1, \frac{N}{2}$$

$N_F$  is the number of frequency channels

$B$  is the bandwidth

In the frequency domain the Hamming taper can be represented by the following smoothing function (e.g. Harris (1978)):

$$H'_i = 0.23U_{i-1} + 0.54U_i + 0.23U_{i+1} \quad (8.9)$$

In a similar way as in the previous section it can be shown that the noise in a Hamming tapered channel is equal to

$$\sigma_{H'_i} = \sqrt{2(0.23)^2 + (0.54)^2} = 0.6303$$

thus the noise equivalent bandwidth will be  $1/(0.63)^2 b = 2.52b$ , where  $b$  is the channel separation,  $B/N_F$ .

The average noise as a result of summing  $n$  adjacent channels will be:

$$\sigma \frac{1}{N} \sqrt{4 * 0.23 * (0.23 - 1) + N} = \frac{1}{N} \sqrt{N - 0.708}$$

## 8.8 REFERENCES

Abramowitz, M. and Stegun, I.A., (1965): "*Handbook of Mathematical functions*" Dover Publications Inc., New York. ISBN 0-486-61272-4.

Bracewell, R.N. (1978): "*The Fourier Transform and Its Applications (2nd edition)*" McGraw-Hill International Book Company. ISBN 0-07-007013-X.

Brigham, E.O. (1974): "*The Fast Fourier Transform*" Prentice-Hall.

Harris, F. (1978): 'On the use of windows for Harmonic analysis with the discrete Fourier Transform' *Proc IEEE*, **66** no.1, pp. 51-83.

More about the measurement of power spectra can be found in:

Blackman, R.B and Tukey, J.W., (1959): "*The measurement of power spectra : from the point of view of communications engineering*". Dover publications, New York. ISBN 0-486-60507-8 first appeared in: Bell system technical journal **37** (1958).

## BIBLIOGRAPHY AND REFERENCES

### 9.1 BOOK REVIEWS

This section contains some reviews of popular textbooks. The number of books on the subject is very large and can of course not all be mentioned. The intention is to give the reader an idea where to start when searching for literature on a particular subject. One may expect that the books discussed can be found in every astronomical library. All necessary information to order the book is cited.

#### 1. **Tools of Radio Astronomy** *K. Rohlfs*

Springer Verlag, Berlin (1986)

ISBN: 3-540-16188-0/0-387-16188-0

This book, which ‘*grew out of a one year graduate course*’, covers a large range of topics. The topics are discussed briefly using a mathematical approach where possible. Although the book is written in a compact style it can be used as a textbook for astronomers/physicist new to the subject of radio astronomy. This book is particularly useful when searching for references on a certain topic; for each chapter it contains general references to standard textbooks/articles.

The book is organized in 13 chapters.

Chap. 1.	Radio Astronomical Fundamentals
Chap. 2.	Electromagnetic Wave Propagation
Chap. 3.	Wave Polarization
Chap. 4.	Fundamentals of Antenna theory
Chap. 5.	Filled Aperture Antennas
Chap. 6.	Interferometers and Aperture Synthesis
Chap. 7.	Receivers
Chap. 8.	Emission Mechanisms of Continuous Radiation
Chap. 9.	Some examples of Thermal and Nonthermal Radio Sources
Chap. 10.	Line Radiation Fundamentals
Chap. 11.	Line Radiation of Neutral Hydrogen
Chap. 12.	Recombination Lines
Chap. 13.	Interstellar Molecules and Their Line Radiation
Appendices	on vector relations, Fourier and Hankel transform, Electromagnetic Field Quantities and Calibration
	Radio Sources

#### 2. **Synthesis Imaging in Radio Astronomy** *Ed. R.A. Perley, F. Schwab, A.H. Bridle*

Astronomical Society of the Pacific (1989)

ISBN: 0-937707-23-6

These are the proceedings of a summer school on Synthesis imaging held in Socorro, New Mexico in June 1988. The course was organized by the National Radio Astronomical Observatory (NRAO) for potential users of the VLA and the VLBI network.

The lectures in the book can be divided in two parts; One part gives an introduction into the theory of aperture synthesis (Chap. 1-11), the other part (Chap 12-25) covers more advanced topics. This book is very useful for beginning and experienced radio astronomers. The topics are all related to aperture synthesis and often related to working with the VLA. This is, however, not a disadvantage because in many respects the WSRT operates similar to the VLA. The topics are discussed in great detail by various authors thus in various styles.

The book contains detailed discussion of aperture synthesis and related topics.

Chap. 1.	Coherence in Radio Astronomy	B.G. Clark
Chap. 2.	The Interferometer in Practice	A.R. Thompson
Chap. 3.	The Primary Antenna Elements	Peter J. Napier
Chap. 4.	Cross Correlators	L.R. D'Addario
Chap. 5.	Calibration and Editing	E.B. Fomalont and R.A. Perley
Chap. 6.	Imaging	R.A. Sramek and F.R. Schwab
Chap. 7.	Sensitivity	P.C. Crane and P.J. Napier
Chap. 8.	Deconvolution	T. Cornwell and R. Braun
Chap. 9.	Self-Calibration	T. Cornwell and E.B. Fomalont
Chap. 10.	Error Recognition	R.D. Ekers
Chap. 11.	Image Analysis	E.B. Fomalont
Chap. 12.	Spectral Problems in Imaging	W.D. Cotton
Chap. 13.	Wide Field Imaging I: Bandwidth and Time-Average Smearing	A.H. Bridle and F.R. Schwab
Chap. 14.	Wide Field Imaging II: Imaging with Non-Coplanar Arrays	R.A. Perley
Chap. 15.	Wide Field Imaging II: Mosaicing	T. Cornwell
Chap. 16.	High Dynamic Range Imaging	R.A. Perley
Chap. 17.	Spectral Line Imaging I: Introduction	P. Roelfsema
Chap. 18.	Spectral Line Imaging II: Calibration and Analysis	J.H. van Gorkum and R.D. Ekers
Chap. 19.	Very Long Baseline Interferometry I: Principles and Practice	R.G. Walker
Chap. 20.	Very Long Baseline Interferometry II: The techniques if Spectral line VLBI	P.J. Diamond
Chap. 21.	Solar Imaging with a Synthesis Telescope	T. Bastian
Chap. 22.	Synthesis Imaging of Spatially Coherent Objects	K.R. Anantharamaiah, T.J. Cornwell and R. Narayan
Chap. 23.	Noise in Imaging of Very Bright Sources	K.R. Anantharamaiah, R.D. Ekers, V. Radhakrishnan, T.J. Cornwell and W. Miller Goss
Chap. 24.	Synthesis Observing Strategies — A 'Hitch-Hikers Guide'	A.H. Bridle
Chap. 25.	The Design of Aperture Synthesis Arrays	R.M. Hjellming

### 3. **Interferometry and Synthesis in Radio Astronomy** *A.R. Thompson, J.M. Moran and G.W. Swenson jr.*

John Wiley & Sons, New York (1986)

ISBN: 0-471-80614-5

The authors of this book, who are involved in operating the VLA, lay emphasis on the technical aspects of radio interferometry. This book is of use for the astronomers who are interested in the theory and implementation of radio techniques, especially interferometry. However, the book is not written from an engineer's point of view. A clear mathematical notation is used throughout the book and mathematics is used to clarify, not to explain.

The book can be used by both the astronomers novice in the field of radio interferometry as the more experienced users. There is a clear division between the introductory sections and the more advanced topics.

The book contains detailed discussion of aperture synthesis and related topics.

Each chapter contains a bibliography and references. It contains 15 Chapters.

Chap. 1.	Introduction and Historical Review
Chap. 2.	Introductory Theory of Interferometers and Correlator Arrays
Chap. 3.	Further Theory of the Interferometer Response
Chap. 4.	Geometrical Relationships and Other Practical considerations
Chap. 5.	Design of Arrays
Chap. 6.	Response of the Receiving System
Chap. 7.	Design of the Analog Receiving System
Chap. 8.	Digital Signal Processing
Chap. 9.	Very-Long-Baseline Interferometry
Chap. 10.	Calibration and Fourier Transformation of Visibility Data
Chap. 11.	Image Processing and Enhancement
Chap. 12.	Interferometer Techniques for Astrometry and Geodesy
Chap. 13.	Propagation Effects
Chap. 14.	Radio Interference
Chap. 15.	Related Techniques

#### 4. **Radio Telescopes** *W.N. Christiansen and J.A. Högbom*

Cambridge University Press (1969, 1985(2<sup>nd</sup> ed.))

ISBN: 0-521-26209-7 (2<sup>nd</sup> ed.) / 0-521-07054-6 (1<sup>st</sup> ed.)

The writers of this book ‘... aim to present to newcomers in radio astronomy a short survey of the development of radio telescopes (...) with enough simplified theory to enable them to understand the fundamentals of radio telescope design.’

In the book there is an emphasis on the relation between aperture distribution, grading, beamshape, etc. A number of possible antenna forms and configurations is discussed, aperture synthesis is only one of antenna systems discussed. The chapter on aperture synthesis has been rewritten completely for the 2<sup>nd</sup> edition.

Chap. 1.	Introduction
Chap. 2.	Some Theory
Chap. 3.	The Steerable Parabolic Reflector (Paraboloid of Revolution)
Chap. 4.	Other Types of Filled-Aperture Antennas
Chap. 5.	Some More Theory
Chap. 6.	Unfilled-Aperture Antennas
Chap. 7.	Synthesis Techniques
Chap. 8.	Sensitivity
App. 1.	Celestial Coordinate System
App. 2.	The Fourier Transform
App. 3.	Available Correlated Power

#### 5. **Galactic and Extragalactic Radio Astronomy (2nd edition)** *Ed. G.L. Verschuur and K.I. Kellerman*

Springer Verlag, Berlin (1988)

ISBN: 0-387-96575-0 / 3-540-96575-0

The second edition of Galactic and Extragalactic Radio Astronomy is ‘intended for graduate students and practicing astronomers who wish to familiarize themselves with the wealth of astronomical phenomena that are “visible” at radio frequencies’.

The book does not discuss the technical aspects of radio astronomy. (In the first edition there was a very good introductory chapter on radio interferometry, but that one has been removed in the second edition). However, a large number of astrophysical and observational topics is discussed. The book gives a broad overview of the kind of astronomy possible using the technique of radio astronomy.

The book has 15 chapters written by different authors. Each chapter contains a section “Recommended Reading” giving reference to standard books and articles in the field.

Chap. 1.	Galactic Nonthermal Continuum Emission	C.J Salter and R.L. Brown
Chap. 2.	HII Regions and Radio Recombination Lines	M.A. Gordon
Chap. 3.	Neutral Hydrogen and the Diffuse Interstellar Medium	S.R. Kulkarni and C. Heiles
Chap. 4.	Molecules as Probes of the Interstellar Medium and of Star Formation	B.E. Turner
Chap. 5.	Interstellar Molecules and Astrochemistry	B.E. Turner and L.M. Ziurys
Chap. 6.	Astronomic Masers	M.J. Reid and J.M. Moran
Chap. 7.	The Structure of Our Galaxy Derived from Observations of Neutral Hydrogen	W. Butler Burton
Chap. 8.	The Galactic Center	H.S. Liszt
Chap. 9.	Radio Stars	R.M. Hjellming
Chap. 10.	Supernova Remnants	S.P.Reynolds
Chap. 11.	Pulsars	D.C. Backer
Chap. 12.	Extragalactic Neutral Hydrogen	R. Giovanelli and M.P. Haynes
Chap. 13.	Radio Galaxies and Quasars	K.I. Kellerman and F.N. Owen.
Chap. 14.	The Microwave Background Radiation	J.M. Uson and D.T. Wilkinson
Chap. 15.	Radio Sources and Cosmology	J.J. Condon

#### 6. **Radio Astronomy (2nd edition)** *J.D. Kraus*

Cygnus-Quasar Books, Ohio (1986)

ISBN: -

‘Radio Astronomy embraces a wide range of topics from astrophysical phenomena to receiver and antenna design. The aim of this book is to bring together a balanced selection and treatment of these topics that is elementary enough to serve as an introduction to radio astronomy yet is sufficiently detailed to be useful as a teaching text and reference work.’

Radio Astronomy is one of the classic text books in the field. The more technical chapters in this book are written from an engineers point of view.

Chap. 1	Introduction
Chap. 2	General Astronomy Fundamentals
Chap. 3	Radio-Astronomy Fundamentals
Chap. 4	Wave Polarization
Chap. 5	Wave-Propagation Fundamentals
Chap. 6	Radio-Telescope Antennas
Chap. 7	Radio-Telescope Receivers <i>by M.E. Tiuri and A.V. Räsänen</i>
Chap. 8	The Radio Sky, Spectra, The Solar System and Our Galaxy
Chap. 9	Pulsars
Chap. 10	Extragalactic Radio Astronomy
Chap. 11	Radio Surveys
Chap. 12	SETI
App. 1	List of Radio Sources
App. 2	Messier's List of Nebulous Objects
App. 3	Frequencies Allocated for Radio Astronomy
App. 4	Relation of Beamwidth and Side-Lobe Level to Aperture Distribution
App. 5	Noise-Temperature-Noise-Figure Chart
App. 6	Precession chart
App. 7	Equatorial- to Galactic-Coordinate Conversion

## 9.2 DICTIONARY

To provide quick reference to the literature and to help the aspirant and unexperienced radio astronomer cope with the jargon we present a dictionary with a couple of the most common words in radio astronomy. We included references to the books from the previous section (e.g. book: No. **1** pp. 6) or to the User Documentation (e.g. userdoc: **II-9-13**) where the concepts are introduced or explained.

**aliasing** — The Fourier transform from the  $u, v$  plane to the image plane will produce a replicated brightness distribution. If the brightness distribution contains structures that are under-sampled in the  $u, v$ -plane then structures in the replicated brightness distribution will overlap and aliasing occurs.

references: book: No. **3** pp. 110-111, userdoc: part II chapter 5.

**antenna pattern** — The antenna pattern, ( $A$ ), also called power pattern, describes the sensitivity of the antenna to power from different directions. Often normalized by dividing through the maximum value of the pattern,  $A_N = A/A_{max} = A/A(0)$ .

references: book: No. **1** pp. 62, book: No. **6** pp. 24.

**bandwidth smearing** — Bandwidth smearing refers to distortion occurring on the edges of maps, far from the fringe stopping center (=phase center), when observing with large bandwidth. It occurs because the phase correction is applied for only one frequency in the band so other frequencies will have slight phase errors, also referred to as chromatic aberration.

references: book: No. **2** pp. 32, 247-253, book: No. **3** pp. 169.

**baseline-pole** — Baseline-pole is defined by the line through the two elements of an interferometer. For the Westerbork array the declination of the baseline pole is defined as the declination of the line through the fixed telescopes. An error in baseline pole introduces a phase error in the visibility data.

references: userdoc: part IV chapter 1.1.

**channel map** — Map of the brightness distribution in one frequency channel

**CLEAN** — Algorithm used to deconvolve maps *i.e.* remove the effects of the dirty beam.

references: userdoc: **II-5-7**, Högbom (1974), book: No. **2** pp. 167-181. book: No. **3** pp. 343-349

**coherence function** — The (source) coherence function is a measure of the coherence of radiation coming from different points of the source. A source is incoherent when radiation coming

from different points on the source is incoherent. The coherence function is then zero.

In many textbooks the theory of aperture synthesis is explained in terms of the coherence function.

references: book: No. **3** pp. 60-63 book: No. **1** pp. 101-105 book: No. **2** pp. 1-9 For synthesis imaging of coherent sources see: book: No. **2** pp. 415-427

**degradation factor** — Factor with which the sensitivity decreases if instead of an analogue correlator a digital correlator is used

**dirty beam** — The dirty beam is the Fourier transform of the visibility coverage *i.e.* all the (complex)  $u, v$  points on which the visibility is measured. In the image plane the dirty beam corresponds to the image of a point source in the phase center.

**dirty map** — Dirty map is the map which is obtained when Fourier transforming the visibility data. It is the convolution of the brightness distribution and the dirty beam.

**fringe** — Oscillating output signal of a correlator correlating the signal from two antennas which form an interferometric element. The frequency of the oscillation is called the fringe rate or fringe frequency. The amplitude of the fringe depends on the power in the spatial frequency to which the interferometer is sensitive. For a point source the phase of the fringe is a measure of the position of the source relative to the phase reference point.

references: book: No. **3** pp. 91-92, book: No. **6** pp. 6,20, userdoc: **II-2-1**, more about fringe frequency for the WSRT: userdoc: **III-10-1**

**fringe stopping** — or fringe rotation reduces the fringe rate to zero by introducing phase rotation and extra delay. This is done to obtain maximum signal over the bandwidth and to reduce the sample frequency with which the fringes need to be measured.

references: book: No. **3** pp. 149 book: No. **2** pp. 19,77

**grading** — or illumination. The grading is a complex function that describes the current distribution over the aperture.

references: userdoc: **II-5-1**, book: No. **4** pp. 29.

grating ring — ellipsoidal structure around sources in a map. They are the result of the radial sampling in the  $u, v$  plane. If the  $u, v$  sampling is done of baseline increments of  $\Delta D_\lambda$  the grating rings will be ellipsoids of semi-axes  $k/\Delta D_\lambda$  and  $k/(\Delta D_\lambda \sin \delta)$  radians in right ascension and declination. Grating rings can be removed using CLEAN.

references: userdoc: **II-5-6** book: No. **1** pp. 117 book: No. **6** pp. 6-33,34

Hamming — Taper function, in the time domain it has the form:

$$f(t) = \begin{cases} 0.54 + 0.46 \cos\left(\frac{\pi t}{T}\right) & \text{for } |t| \leq T \\ 0 & \text{for } |t| \geq T \end{cases}$$

This function goes smoothly to zero at time  $T$ .

userdoc: part II chapter 8. book: No. **3** pp. 339

Hanning — Taper function named after van Hann, an Austrian mathematician. In the time domain it has the form:

$$f(t) = \begin{cases} \frac{1}{2} \left(1 + \cos\left(\frac{\pi t}{T}\right)\right) & \text{for } |t| \leq T \\ 0 & \text{for } |t| \geq T \end{cases}$$

This function goes smoothly to zero at time  $T$ .

The Fourier transform of this window function in the frequency domain, or the observed spectral response at frequency  $\nu_0$  to a monochromatic point source situated at frequency  $\nu$ , is given by

$$F(\nu_0 - \nu) = \frac{1}{2}S(\nu_0 - \nu) + \frac{1}{4}S(\nu_0 - \nu + b) + \frac{1}{4}S(\nu_0 - \nu - b)$$

where  $b = B/N_F$  is the channel separation. ( $B$  is total bandwidth and  $N_F$  is the number of frequency channels)

references: userdoc: part II chapter 8. book: No. **3** pp. 339

Jansky — flux density unit named after radio astronomy pioneer Karl G. Jansky who was the first to measure radio radiation from the sun in 1932.  $1 \text{ Jy} = 10^{-26} \text{ Watts m}^{-2} \text{ Hz}^{-1}$

image domain — or image plane contains a description the brightness distribution as function of some coordinate on the sky.

mosaicing — Process of combining multiple observations with different pointings into one single map with a field of view that may be larger than the field of view of a single primary beam.

references: userdoc: part III chapter 6 book: No. **2** pp. 277-286

phase reference point — also fringe stopping point, or phase center. Position on the sky for which, for all observing frequencies, the phase of the visibility function is zero. In most cases the center of the field *i.e.* the pointing position is chosen as such.

references: userdoc: part II chapter 2 book: No. **2** pp. 14 book: No. **3** pp. 80

polarization angle — Angle of linear polarized radiation as measured from north ( $0^\circ$ ) through east ( $90^\circ$ ).  $PA = 1/2 \arctan(U/Q)$  where  $U$  and  $Q$  are Stokes parameters.

references: userdoc: **III-4-1**

primary beam — In principle this is the antenna response power pattern of a single antenna element. In practice the primary beam also corrects for other instrumental effects which cause signal attenuation.

references: userdoc: part III, section 8.1.

sensitivity — is a measure of the weakest source which, in the absence of confusing sources, can be detected with confidence. It is often defined as the signal corresponding to the r.m.s. error deflections in a map due to thermal noise.

references: userdoc: **II-7-3** book: No. **2** pp. 139-165 book: No. **3** pp. 155-168 book: No. **4** pp. 226-246 book: No. **6** pp. 7-10

Shah function — Function introduced by Bracewell (e.g. 1978) also called sampling function. It is defined by

$$\text{III}(x) = \sum_{n=-\infty}^{+\infty} \delta(x - n)$$

references:

Bracewell, R.N. (1978): "*The Fourier Transform and Its Applications (2nd edition)*" McGraw-Hill International Book Company. ISBN 0-07-007013-X.

userdoc: **II-4-1**

Stokes parameters — The Stokes parameters ( $I, Q, U,$  and  $V$ ) are used to describe the polarization properties of electromagnetic waves.  $I^2 = Q^2 + U^2 + V^2$  is a measure for the total intensity in a wave.  $U$  and  $Q$  are a measure for the amount of linear polarization and  $V$  is a measure for the amount of circular polarization. If  $V = 0$  there is only linear polarization. If  $U = Q = 0$  then the polarization is circular.

references: userdoc: part III, chapter 4. book: No. **1** pp. 36-39

temperature, antenna — temperature related to the power received by a radio antenna.  $T_A = \frac{W_a}{\Delta\nu k}$  where  $W_a$  is the power received in a frequency band with width  $\Delta\nu$  and  $k$  is the Boltzmann's constant.

references: userdoc: part II chapter 7 book: No. 1 pp. 68 book: No. 3 pp. 10

temperature, brightness — The temperature related to the power emitted by a source (by means of the Rayleigh-Jeans law)

references: userdoc: part II chapter 7 book: No. 1 pp. 9-13 book: No. 3 pp. 8

temperature, system — Temperature corresponding to the noise power of the receiver and all the other sources of unwanted noise. Important for the determination of sensitivity.

references: userdoc: **II-7-4** book: No. 2 pp. 139-165 book: No. 1 pp. 133 book: No. 3 pp. 155-168 book: No. 6 pp. 7-10

taper — or window functions are used to control the shape of a beam after Fourier transformation. They are used when in two cases:

- the cross correlation function (CFF) is tapered to reduce the sidelobes of the beam in the frequency domain.
- the visibilities in the  $u, v$  plane are tapered to reduce the sidelobes of the beam in the image plane

In the spatial frequency plane (*i.e.*  $u, v$ -plane) the grading function is often used to taper the data. (The grading actually is the combination of a taper and a  $u, v$ -sampling function.)

references: userdoc: part II chapter 5 and 8

tied array — in the tied array mode all telescopes of the array are autocorrelated to obtain a large collective area. High spatial resolution is not obtained. This mode is used for VLBI observations when the WSRT is one station in the array.

references: userdoc: **III-1-3** book: No. 3 pp. 339

VLBI — Very Long Baseline Interferometry. In Westerbork two special backends/recorders are available for participating in VLBI observations.

references: userdoc: **III-1-3** book: No. 2 pp. 355-393

visibility — or visibility function describes the brightness distribution in terms of its Fourier components (the fringes). (Its units are  $\text{W m}^{-2}\text{Hz}^{-1}$ .) Thompson (book: No. 2 pp. 4) describes the visibility as used in astronomy 'visibility is a complex quantity, the magnitude of which has the dimension of spectral power flux density. It can

be regarded as an unnormalized measure of the coherence of the electric field, modified to some extent by the characteristics of the interferometer.'

references: book: No. 3 pp. 55 book: No. 1 pp. 109 book: No. 6 pp. 6-24

Westerbork Unit — equals 5mJy, unit used for the standard WSRT visibility output on disk/tape.