



Digital Systems for the MITRA

(GPU Computing)

Submitted by

Ruben Anderson Louis

Project submitted in the partial fulfillment for the degree of

BSc (Hons) Physics with Computing

UNIVERSITY OF MAURITIUS

FACULTY OF SCIENCE

DEPARTMENT OF PHYSICS

March, 2015

Declaration of Authorship

I, AUTHOR NAME, declare that this thesis titled, 'THESIS TITLE' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

“Write a funny quote here.”

If the quote is taken from someone, their name goes here

UNIVERSITY OF MAURITIUS

Abstract

Faculty of Science

Department of Physics

BSc (Hons) Physics with computing

Digital Systems for the MITRA

by [Ruben Anderson Louis](#)

In this report we give a brief account about imaging in the field of Radio Astronomy specifically on the technique coined Aperture Synthesis.

Acknowledgements

I am grateful to the Dr. G. K. Beeharry for his useful suggestions, guidance and remarks, and for the opportunity he gave me to work on the topic. Also I express my thanks to the following for their help: Assoc. Prof. R. Somanah, Dr S. Oree and the technical staff of the department of physics.

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iv
List of Figures	vii
List of Tables	viii
Abbreviations	ix
Physical Constants	x
Symbols	xi
1 Introduction	1
1.1 About the report	1
1.2 The pioneers	2
1.3 Introductory basic concepts	2
1.3.1 Imaging by Interferometry in R.A.	2
1.3.2 Why are radio interferometers used?	3
1.3.3 The principle of a radio Interferometer	5
1.3.4 A free coordinate system	6
1.3.5 Visibility – Sky intensity distribution relationship	8
2 Literature Review	10
3 Cross-correlation	19
3.1 The correlator	19
3.2 Cross-correlation	20
4 Image Reconstruction	23
4.1 The Van Cittert-Zernike theorem	23
4.2 Sampling of the visibility	24
4.2.1 Sampling Theorem	25

4.2.2	Array configuration consideration	28
4.2.2.1	Spatial Sensitivity and the Spatial Transfer function	29
4.3	Data calibration	34
4.3.1	Calibration using calibration sources	34
4.4	Image Reconstruction	36
4.4.1	Weighting of the visibility	37
4.4.2	Mapping by Discrete Fourier transform	37
5	Image Processing	39
5.1	Introduction - Image Processing	39
5.2	Deconvolution Procedures	41
5.2.1	The CLEAN algorithm	41
5.2.2	Performance of the CLEAN algorithm	45
6	Conclusions on the report	48
A	Appendix Title Here	49
	Bibliography	50

List of Figures

1.1	Sir Martin Ryle	2
1.2	Quasar 3C273	2
1.3	Diagrammatic representation of an elementary radio interferometer setup	2
1.4	Cygnus A	3
1.5	Diffraction Effect	3
1.6	Antenna Pattern	4
1.7	Angular Resolution	4
1.8	An elementary interferometer	5
1.9	Free coordinate system	6
1.10	Element of Solid Angle	7
3.1	A simplistic representation of a correlator	19
4.1	Fringe Visibility	23
4.2	Amplitude and fringe of the visibility function	24
4.3	Amplitude and fringe of the visibility function, where the source is offset the phase centre	24
4.4	Sampling a 1-D visibility	25
4.5	Sampling a 2-D visibility	27
4.6	A correlator array	28
4.7	Antenna aperture configurations and autocorrelation nonzero boundaries	29
4.8	Visibility map coverage and derived Intensity map	33
4.9	The effect of weighting visibilities	37
5.1	The end result of image reconstruction	39
5.2	CLEAN process	44
5.3	Point-source response (broken line) removal at the maximum of a broad feature, Clark (1982)	46

List of Tables

Abbreviations

EM	Electromagnetic
FFT	Fast Fourier Transform
MEM	Maximum Entropy Method
MRT	Mauritius Radio Telescope
NNLS	non-negative least square algorithm
R.A.	Radio Astronomy
VLA	Very Large Array (Radiotelescope)

Physical Constants

Speed of Light $c = 2.997\ 924\ 58 \times 10^8\ \text{ms}^{-1}$ (exact)

Symbols

a	distance	m
P	power	W (Js ⁻¹)
ω	angular frequency	rads ⁻¹

For/Dedicated to/To my...

Chapter 1

Introduction

1.1 About the report

The aim of this report is to introduce the topic of Imaging in Radio Astronomy to students with an introductory university background in signal & image processing and having a secondary school knowledge of physics. This was done so that the reader would acquire the basic knowledge about what is done actually, the practical issues, and thus those interested to learn or try out things on their own on the topic would be encouraged to read the literature referred to.

The report is composed of 4 main chapters where the literature from the book, **Interferometry and Synthesis in Radio Astronomy** by Thompson, Moran, and Swenson Jr [6] has been put and reformulated in a concise way along with literature from other sources, so original work is not expected here. The report is structured as follows *chapter 1* is an introduction to the field, the main goals people aim for in the area and it also stands as a basis for the other chapters. Then *chapter 3* continues more thoroughly on the topic of cross-correlation, and we pass on to *chapter 4* which deals more with theorems, sampling considerations, data calibration, and image reconstruction/retrieval, generally things related to linear processes. Then finally *chapter 5* expounds about the actual processing of the raw image with enhancing techniques which usually deal with non-linear processes. So we hope that the reader will find this small piece of work that we put together very fluent and comprehensive on the topic. Feel free to mail us on our project group mail the-radio-imagists@googlegroups.com if you have questions on the topic discussed in the report and/or suggestions on the report itself.

1.2 The pioneers

As Assoc. Prof. R. Somanah usually says, one cannot talk about Radio Imaging without mentioning a main pioneer in the field, Sir Martin Ryle. He was a British radio astronomer who did develop revolutionary radio telescopes systems and used them for the accurate location and imaging of weak radio sources. Earlier he worked on the study of radio waves from the Sun and sunspots, and later discovered the first quasi-stellar object known as the Quasar². Martin Ryle and his colleague Anthony Hewish were the first **astronomers** to ever receive the Nobel prize in Physics in 1974 for their overall contribution to **radio astronomy** (Encyclopedia Britannica [2]).



FIGURE 1.1:
Sir Martin Ryle¹

1.3 Introductory basic concepts

1.3.1 Imaging by Interferometry in R.A.

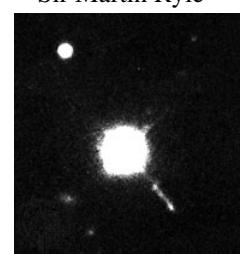
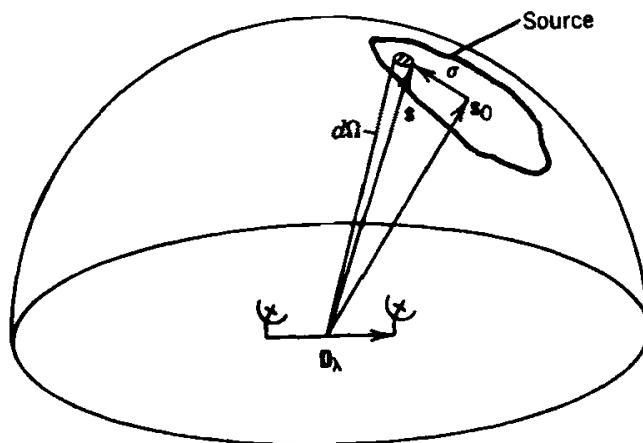


FIGURE 1.2: Quasar
3C273 [3]

FIGURE 1.3:
A Radio interferometer setup diagrammatic representation [6, Pg. 69, Fig. 3.1]

Interferometry is a technique where electromagnetic waves are superimposed in order to extract information about the waves. As concern Radio Astronomy when at least 2 radio telescopes are working in tandem, the setup is given the name of radio interferometer. Radio Telescopes are actually antennas which basically are devices that respond to incoming electromagnetic radiation and output electrical currents related to this response. For application in Radio Astronomy passive radio telescopes i.e. receiving antennas are used.

¹Stamp, Sir Martin Ryle - Radio Surveyor of the Universe.
http://colnect.com/en/stamps/stamp/184650-Sir_Martin_Ryle_-_Radio_Surveyor_of_the_Universe-Eminent_Britons-United_Kingdom_of_Great_Britain_Northern_Ireland

²Quasar, an astronomical object of very high luminosity found in the centres of some galaxies and powered by gas spiraling at high velocity into an extremely large black hole. Quasar 3C 273, the brightest and closest of the quasi-stellar radio sources (Encyclopedia Britannica [3]).

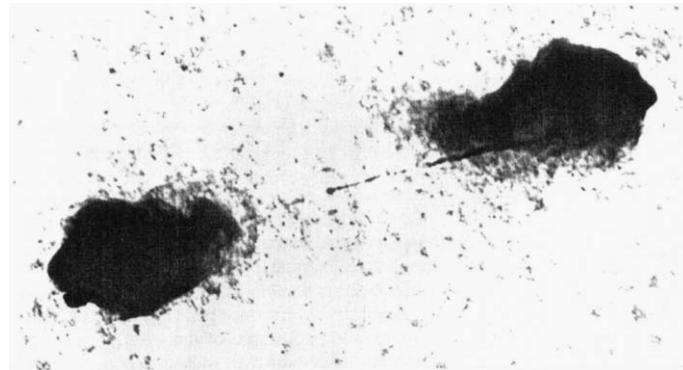


FIGURE 1.4:
Radio image of Cygnus A made with the VLA at 4.9 GHz by Perley, Dreher, and Cowan (1984) [6,
Pg. 33, Fig. 1.18]

1.3.2 Why are radio interferometers used?



FIGURE 1.5:
Diffraction Effect [8, Appendix B.1, Fig. B.1]

(From Woods et al. [8, Appendix B.1]) To answer this question one must introduce the concept of angular resolution, the latter describes the angular distance between two point sources that can be differentiated by an aperture. Because of the diffraction effect, an antenna pattern (Fig. 1.6) has side lobes, which are sensitive to sources outside the main antenna beam, limiting resolution. When a planar electromagnetic wave enters an aperture, the electromagnetic wave is distorted in what is called a diffraction pattern (Fig. 1.5). Therefore a finite sized aperture cannot correctly record the radio brightness without some distortion of the original signal. The diffraction distortion is due to the interaction of the original EM wave with the edges of a finite sized aperture, which creates the fringe pattern of destructive and constructive interference. Diffraction affects all types of EM waves when entering an aperture, but is more severe for longer wavelengths. The distance to the first zero of the diffraction pattern of a circular aperture is given by

$$R \simeq 1.22 \frac{\lambda}{D} \quad (1.1)$$

If two objects are closer than the first minima (1.7(a)), for a particular aperture, they cannot be distinguished. Therefore the first minima, determines the resolving capabilities of an aperture and

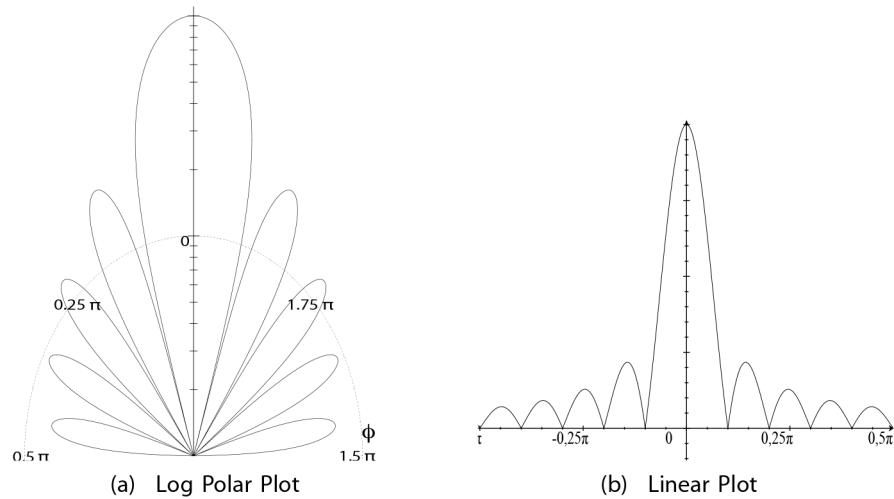


FIGURE 1.6:
Antenna Pattern [8, Appendix B.1, Fig. B.2]

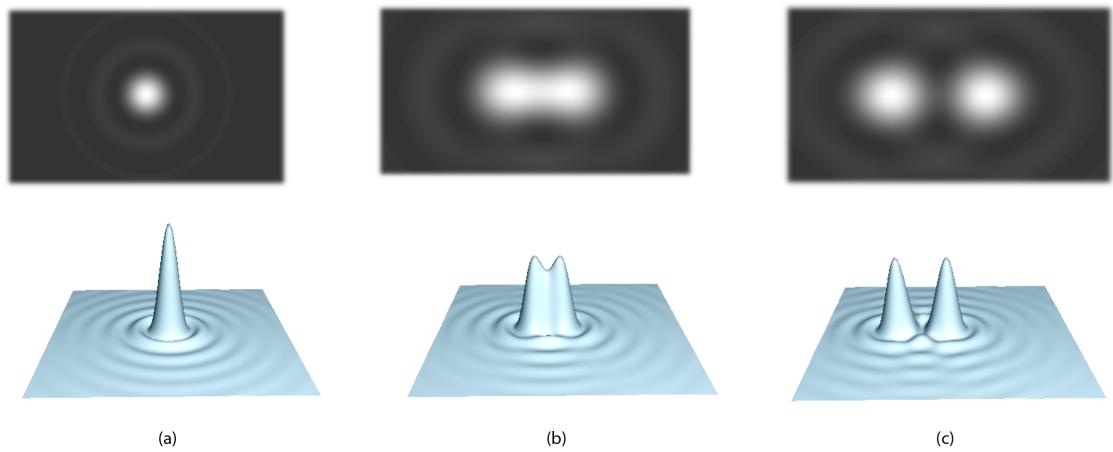


FIGURE 1.7:
Angular Resolution [8, Appendix B.1, Fig. B.3]

is called the angular resolution R . The angular resolution, represented in the right-hand side of the equation 1.1, depends on both the wavelength and the aperture diameter. As a consequence of dealing with radio waves, which have a long wavelength, radio astronomy requires large telescopes in order to improve the resolution and produce detailed images (radio brightness readings). The dimensions of a single radio aperture needed to meet the angular resolution requirements are extremely impractical in terms of strict design requirements and physical constraints. For example, to achieve the same angular resolution as the naked human eye, a radio antenna's aperture observing a source at 4GHz must be 750m in diameter. Therefore the way this issue is coped for is by using a radio interferometer setup, in the latter case its effective aperture diameter now corresponds to the largest separation of the 2 telescopes and actually a much better resolution can be obtained more easily, as one can have telescopes with a desired distance between them (in that

small calculation $\sim 750\text{m}$) which is more practical than having to build a single radio telescope with a large aperture diameter. This is from that fact that the technique is coined *Aperture Synthesis* as with that setup the aperture diameter and in consequence the resolution of a much larger telescope can be emulated by the aforementioned mean. Knowing the impulse response of an aperture, a closer reconstruction of the original source can be made by performing a deconvolution, this will be discussed in *chapter 5*

1.3.3 The principle of a radio Interferometer

[From 6, Sec. 2.1] Consider the figure 1.8 below which shows the direction of incoming electromagnetic planar wavefronts received at the antenna from the sky. Depending on the direction of

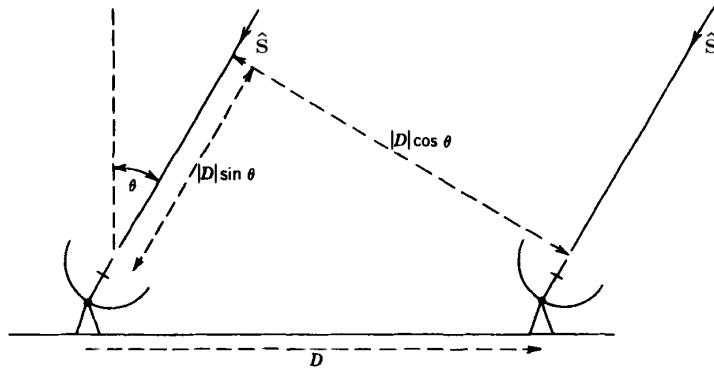


FIGURE 1.8:
Geometry of an elementary interferometer [6, Pg. 51, Fig. 2.1]

the wavefront relative to the direction the antennas point to, here characterised by the angle θ , the same wavefront from the sky will reach a particular antenna first, here the right one, or both would get the same wavefront at the same time, in the first case it is only after a certain lapse of time that the wavefront reaches the other antenna, the left one. The extra distance the wavefront has to travel to reach the left antenna is the projection of the baseline direction on the unit vector in the source direction which is equal to $\mathbf{D} \cdot \hat{\mathbf{s}}$.

$$\mathbf{D} \cdot \hat{\mathbf{s}} = |\mathbf{D}| \cdot \cos(90^\circ - \theta) = |\mathbf{D}| \sin(\theta) \quad (1.2)$$

This is also the extra time that the wavefront takes at speed of light to reach the left antenna known as the geometric delay or time delay. The geometric delay, τ_g is thus the following,

$$\tau_g = \frac{\mathbf{D} \cdot \hat{\mathbf{s}}}{c} \quad (1.3)$$

where, c , is the speed of light in vacuum.

As mentioned in section 1.3.1 antennas output electrical signals in response to the incoming wavefront, thus it is quite obvious that the signal output at the left antenna would be similar to that from

the the right antenna but delayed by the geometric delay. Therefore, what we do in radio interferometry is that we have a measure to characterise this antenna pair - signal relationship in response to the EM radiation from a direction in the sky and it is called the **visibility**. The relationship between the visibility and what is observed is direct. So one can make a map of the visibility and be able to have information about the sky observed and this is discussed further in the section 1.3.5

1.3.4 A free coordinate system

[From 6, Sec. 3.1, Pgs. 68-71] Before we continue further let us first use an adequate coordinate system that will be the basis of our subsequent discussion. Consider the figure 1.9. Suppose that

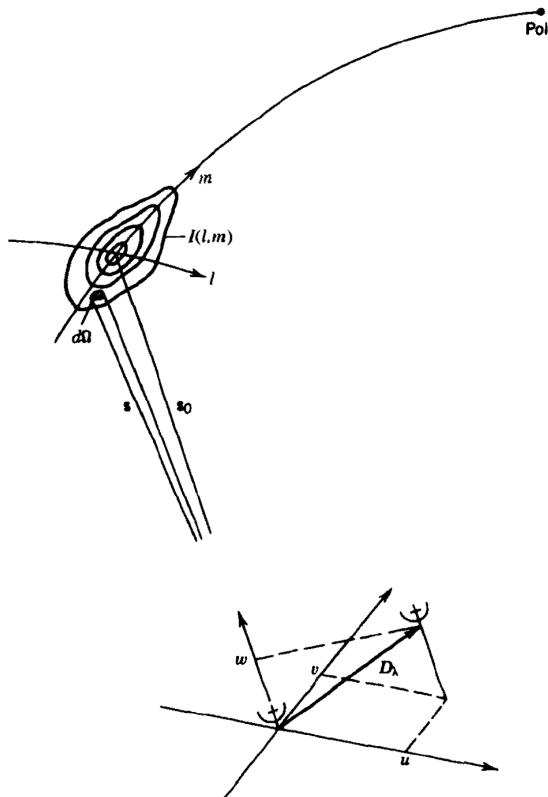


FIGURE 1.9:
Free coordinate system [6, Pg. 70, Fig. 3.2]

the antennas track the source under observation, which is the most common situation, and let the unit vector \hat{s}_0 indicate the phase reference position. This position, known as the phase-tracking centre, becomes the centre of the field to be mapped. The magnitude of the baseline vector, $|D_\lambda| = \frac{|D|}{\lambda}$, is measured in wavelengths at the centre frequency of the observing band, and the baseline direction, D_λ , has components (u, v, w) in a right-handed coordinate system, where u and v are measured in a plane normal to the direction of the phase reference position. The spacing component u is measured toward the north as defined by the plane through the origin, the source, and the pole, and v toward the east. The component w is measured in the direction \hat{s}_0 and so is

defined as follows,

$$\mathbf{D}_\lambda \cdot \hat{\mathbf{s}}_0 = w \quad (1.4)$$

Source direction or position on the celestial sphere have the components (l, m, n) which are simply, respectively the direction cosines of the components of the particular direction/position $\hat{\mathbf{s}}$ in terms of the (u, v, w) components i.e.

$$\hat{\mathbf{s}} = (l, m, n) = (\hat{\mathbf{s}} \cdot \hat{\mathbf{u}}, \hat{\mathbf{s}} \cdot \hat{\mathbf{v}}, \hat{\mathbf{s}} \cdot \hat{\mathbf{w}}) \quad (1.5)$$

Then, since $l^2 + m^2 + n^2 = 1$ we can re-express n in terms of l and m as follows,

$$n = \sqrt{1 - l^2 - m^2} \quad (1.6)$$

leading us to re-express source directions with the following components i.e.

$$\hat{\mathbf{s}} = (l, m, \sqrt{1 - l^2 - m^2}) \quad (1.7)$$

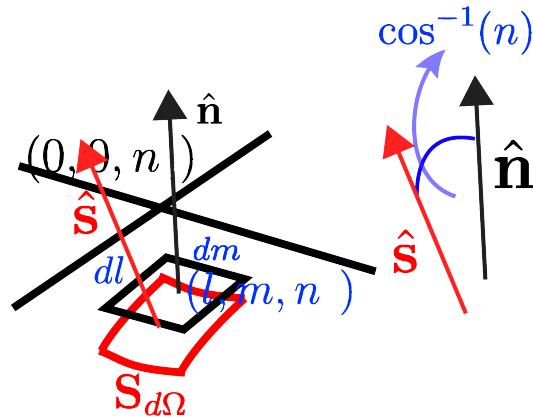


FIGURE 1.10:
Element of Solid Angle

One can also easily show that for an elementary patch of the sky $dl \cdot dm$, the solid angle, $d\Omega$, subtended from our reference is,

$$d\Omega = \frac{dl dm}{n} \quad (1.8)$$

Refer to figure 1.10, note that the solid angle, $d\Omega$ is just the projection of the patch $dl \cdot dm$ on the unit sphere centred on the origin. Referring to the figure we denote the elementary surface at coordinate $(u = l, v = m, n = n)$ as $\mathbf{S}_{dldm} = S_{dldm} \cdot \hat{\mathbf{n}}$ and its projection on the unit sphere as, $\mathbf{S}_{d\Omega} = S_{d\Omega} \cdot \hat{\mathbf{s}}$. So we have the following relationship,

$$\mathbf{S}_{d\Omega} \cdot \hat{\mathbf{n}} = S_{dldm} \quad (1.9)$$

$$S_{d\Omega} \cdot \hat{\mathbf{s}} \cdot \hat{\mathbf{n}} = S_{dldm} \quad (1.10)$$

As the angle between $\hat{\mathbf{s}}$ and $\hat{\mathbf{n}}$ is $\cos^{-1}(n)$

$$S_{d\Omega} \cdot \cos(\cos^{-1}(n)) = S_{dldm} \quad (1.11)$$

$$\begin{aligned} S_{d\Omega} \cdot n &= S_{dldm} \\ S_{d\Omega} &= \frac{S_{dldm}}{n} \end{aligned} \quad (1.12)$$

$$d\Omega = \frac{dldm}{n}$$

$$d\Omega = \frac{dldm}{\sqrt{1 - l^2 - m^2}} \quad (1.13)$$

1.3.5 Visibility – Sky intensity distribution relationship

[From 6, Sec. 3.1, Pg. 71-73] As introduced very roughly in section 1.3.3 there is a direct relationship between the electrical signal received at the pair of antennas and the electromagnetic waves from the sky, we also mentioned the visibility which relates the antenna-pair electrical signal to the electromagnetic waves received. Most importantly the visibility is a function of the baseline vector and relates to the sky intensity distribution in the following way,

$$V'(u, v, w) = \iint_{\text{sky patch}} A_N I(l, m) e^{-j2\pi(\mathbf{D}_\lambda \cdot \hat{\mathbf{s}})} d\Omega \quad (1.14)$$

$$V'(u, v, w) = \iint_{\text{sky patch}} \frac{A_N I(l, m) e^{-j2\pi(uI + vM + w(\sqrt{1-l^2-m^2}-1))}}{\sqrt{1-l^2-m^2}} dldm \quad (1.15)$$

where, $V'(u, v, w)$ is the measured visibility, A_N , the normalised product of the antenna beams, and $I(l, m)$, is the intensity distribution.

Based on an approximation that is valid so long as the synthesised field is not too large. If l and m are small enough that the term

$$\left(\sqrt{(1 - l^2 - m^2)} - 1 \right) w \simeq -\frac{1}{2}(l^2 + m^2)w \quad (1.16)$$

can be neglected then what is left is the following:

$$V'(u, v) = \iint_{\text{sky patch}} \frac{A_N I(l, m) e^{-j2\pi(ul+vm)}}{\sqrt{1 - l^2 - m^2}} dl dm \quad (1.17)$$

which is a direct Fourier transform relationship between visibility map and sky intensity distribution multiplied by the normalised antenna pattern. The actual calculation done to obtain a visibility is discussed in the chapter 3 and we shall focus on the use of a correlator.

Chapter 2

Literature Review

Preliminary gather up of abstract or similar

2008

Jheengut - Software Correlation

Software correlation is seen to replace digital correlation as a step forward in removing the excessive cost for dedicated hardware in the near future. A software correlator for radio astronomy has been designed in the FORTRAN programming language with design considerations as an off the shelf project. The flexibility of a software correlation is being taken as an aid to solve problems found in dedicated hardware to be upgraded and maintained properly.

S. Ord, L. Greenhill, R. Wayth, D. Mitchell, K. Dale, H. Pfister, R. G. Edgar - GPUs for data processing in the MWA

The MWA is a next-generation radio interferometer under construction in remote Western Australia. The data rate from the correlator makes storing the raw data infeasible, so the data must be processed in real-time. The processing task is of order 10 TFLOPs^{-1} . The remote location of the MWA limits the power that can be allocated to computing. We describe the design and implementation of elements of the MWA real-time data processing system which leverage the computing abilities of modern graphics processing units (GPUs). The matrix algebra and texture mapping capabilities of GPUs are well suited to the majority of tasks involved in real-time calibration and imaging. Considerable performance advantages over a conventional CPU-based reference implementation are obtained.

Chris Harris Karen Haines Lister Staveley-Smith -
GPU Accelerated Radio Astronomy Signal Convolution

The increasing array size of radio astronomy interferometers is causing the associated computation to scale quadratically with the number of array signals. Consequently, efficient usage of alternate processing architectures should be explored in order to meet this computational challenge. Affordable parallel processors have been made available to the general scientific community in the form of the commodity graphics card. This work investigates the use of the Graphics Processing Unit (GPU) in the parallelisation of the combined conjugate multiply and accumulation stage of a correlator for a radio astronomy array. Using NVIDIA's Compute Unified Device Architecture, our testing shows processing speeds from one to two orders of magnitude faster than a Central Processing Unit (CPU) approach.

Andrew Woods, Michael Inggs and Alan Langman -

Accelerating a Software Radio Astronomy Correlator using FPGA co-processors

This article presents and characterises our work on accelerating a software radio astronomy correlator using reconfigurable computing (RC) hardware. Radio astronomy correlation is an embarrassingly parallel signal processing application, which is used heavily in radio astronomy for imaging and other astronomical measurements. Radio astronomy correlators typically operate on huge data sets and often require real-time processing, as storage of raw data is impractical - resulting in substantial computational requirement. Currently FPGAs are the preferred processing architecture used in modern large radio astronomy correlators [1] and perform well on the types of DSP functions that correlators perform. In this paper we set out to accelerate the DiFX (Distributed FX) correlator, a software correlator, using FPGA reconfigurable computing hardware. — hoping to inherit some of the advantages that larger production FPGA correlators have over software.

2009

NVIDIA -

NVIDIA's Next Generation CUDA™ Compute Architecture: Fermi™

Rob V. van Nieuwpoort, John W. Romein -

Using Many-Core Hardware to Correlate Radio Astronomy Signals

A recent development in radio astronomy is to replace traditional dishes with many small antennas. The signals are combined to form one large, virtual telescope. The enormous data streams are crosscorrelated to filter out noise. This is especially challenging, since the computational demands grow quadratically with the number of data streams. Moreover, the correlator is not only computationally intensive, but also very I/O intensive. The LOFAR telescope, for instance, will produce over 100 terabytes per day. The future SKA telescope will even require in the order of exaflops, and

petabits/s of I/O. A recent trend is to correlate in software instead of dedicated hardware. This is done to increase flexibility and to reduce development efforts. Examples include e-VLBI and LOFAR. In this paper, we evaluate the correlator algorithm on multi-core CPUs and many-core architectures, such as NVIDIA and ATI GPUs, and the Cell/B.E. The correlator is a streaming, real-time application, and is much more I/O intensive than applications that are typically implemented on many-core hardware today. We compare with the LOFAR production correlator on an IBM Blue Gene/P supercomputer. We investigate performance, power efficiency, and programmability. We identify several important architectural problems which cause architectures to perform suboptimally. Our findings are applicable to data-intensive applications in general. The results show that the processing power and memory bandwidth of current GPUs are highly imbalanced for correlation purposes. While the production correlator on the Blue Gene/P achieves a superb 96% of the theoretical peak performance, this is only 14% on ATI GPUs, and 26% on NVIDIA GPUs. The Cell/B.E. processor, in contrast, achieves an excellent 92%. We found that the Cell/B.E. is also the most energy-efficient solution, it runs the correlator 5-7 times more energy efficiently than the Blue Gene/P. The research presented is an important pathfinder for next-generation telescopes.

2010

Andrew Woods, Michael Inggs and Alan Langman -

Accelerating a Software Radio Astronomy Correlator using FPGA co-processors

This thesis attempts to accelerate compute intensive sections of a frequency domain radio astronomy correlator using dedicated co-processors. Two co-processor implementations were made independently with one using reconfigurable hardware (Xilinx Virtex 4LX100) and the other uses a graphics processor (Nvidia 9800GT). The objective of a radio astronomy correlator is to compute the complex valued correlation products for each baseline which can be used to reconstruct the sky's radio brightness distribution. Radio astronomy correlators have huge computation demands and this dissertation focuses on the computational aspects of correlation, concentrating on the X-engine stage of the correlator. Although correlation is an extremely compute intensive process, it does not necessarily require custom hardware. This is especially true for older correlators or VLBI experiments, where the processing and I/O requirements can be satisfied by commodity processors in software. Discrete software co-processors like GPUs and FPGAs are an attractive option to accelerate software correlation, potentially offering better FLOPS/watt and FLOPS/\$ performance. In this dissertation we describe the acceleration of the X-engine stage of a correlator on a CUDA GPU and an FPGA. We compare the co-processors' performance with a CPU software correlator implementation in a range of different benchmarks. Speedups of 7x and 12.5x were achieved on the FPGA and GPU correlator implementations respectively. Although both implementations achieved speedups and better power utilisation than the CPU implementation, the GPU implementation produced better performance in a shorter development time than the FPGA. The FPGA implementation was hampered by the development tools and the slow PCI-X bus, which is used to communicate with the host. Additionally, the Virtex 4 LX100 FPGA was

released two years before the Nvidia G80 GPU and so is more behind the current technologies. However, the FPGA does have an advantage in terms of power efficiency, but power consumption is only a concern for large compute clusters. We found that using GPUs was the better option to accelerate small-scale software X-engine correlation than the Virtex 4 FPGA.

Nicolas PLATEL - Implémentation d'un corrélateur sur une carte GPU.

Le but de ce projet est d'implémenter un corrélateur de type FX en Software sur une carte GPU de la marque NVIDIA. Ce corrélateur permettra aux étudiants de finir la construction du télescope étudié et obtenir des images du ciel. Il a pour but également de donner quelques notions sur Cuda aux étudiants le désirants. Dans le but d'étudier des phénomènes physiques connus et de valider mon projet, des résultats expérimentaux ont été effectués grâce aux antennes et au récepteur numérique créé précédemment. Enfin, une interface graphique a été créée pour faciliter l'utilisation du corrélateur à l'utilisateur.

Hobiger T., Kimura M., Takefuji K, Oyama T., Koyama Y., Kondo T., Gotoh T., Amagai J.

GPU based software correlators-perspectives for VLBI2010

Caused by historical separation and driven by the requirements of the PC gaming industry, Graphics Processing Units (GPUs) have evolved to massive parallel processing systems which entered the area of non-graphic related applications. Although a single processing core on the GPU is much slower and provides less functionality than its counterpart on the CPU, the huge number of these small processing entities outperforms the classical processors when the application can be parallelized. Thus, in recent years various radio astronomical projects have started to make use of this technology either to realize the correlator on this platform or to establish the post-processing pipeline with GPUs. Therefore, the feasibility of GPUs as a choice for a VLBI correlator is being investigated, including pros and cons of this technology. Additionally, a GPU based software correlator will be reviewed with respect to energy consumption/GFlop/sec and cost/GFlop/sec.

Patrick Brandt, Ron Duplain, Paul Demorest, Randy McCullough, Scott Ransom, Jason Ray

Heterogeneous real-time computing in radio astronomy

Modern computer architectures suited for general purpose computing are often not the best choice for either I/O-bound or compute-bound problems. Sometimes the best choice is not to choose a single architecture, but to take advantage of the best characteristics of different computer architectures to solve your problems. This paper examines the tradeoffs between using computer systems based on the ubiquitous X86 Central Processing Units (CPU's), Field Programmable Gate Array (FPGA) based signal processors, and Graphical Processing Units (GPU's). We will show how a heterogeneous system can be produced that blends the best of each of these technologies into a

real-time signal processing system. FPGA's tightly coupled to analog-to-digital converters connect the instrument to the telescope and supply the first level of computing to the system. These FPGA's are coupled to other FPGA's to continue to provide highly efficient processing power. Data is then packaged up and shipped over fast networks to a cluster of general purpose computers equipped with GPU's, which are used for floating-point intensive computation. Finally, the data is handled by the CPU and written to disk, or further processed. Each of the elements in the system has been chosen for its specific characteristics and the role it can play in creating a system that does the most for the least, in terms of power, space, and money.

GINOURIE Sabera Bibi

A prototype front-end and back-end receiver system for radioastronomy

The first part of the project consisted of designing and building a front-end and back-end system for radioastronomy. Eight Log-periodic dipole antennas (available at the MRT) were used for the front-end. In the second part, a new data acquisition card was used. This card was studied and programmed before used. The card was tested several times in order to check whether the analog data were digitised. Next, the whole system was tested and observations were carried out. Celestial objects like Virgo A and Centaurus A were successfully observed.

2011

V. K. Veligatla, P. Labropoulos, L. V. E. Koopmans -

Adaptive Beam-forming for Radio Astronomy On GPU

The LOFAR radio telescope consists of tens of thousands of dipole antennas that combine their signals to operate as a single large radio telescope. The truly innovative aspect of this new telescope is that its pointing system is not mechanical. It is steered by combining the electric signals from different elements using advanced beam-forming software. Imaging software is one of the important aspects of processing the high-volume data streams produced by LOFAR, and is one of the best places to use GPUs to achieve processing speed. We were able to achieve up to 30 times performance gain compared to the CPU implementation in novel, computationally intensive techniques such as the Minimum Variance Distortionless Response (MVDR). We have gained 5-6 times speed-up compared to the CPU implementation for standard imaging algorithms.

M. A. Clark, P. C. La Plante, L. J. Greenhill -

Accelerating Radio Astronomy Cross-Correlation with Graphics Processing Units

We present a highly parallel implementation of the cross-correlation of time-series data using graphics processing units (GPUs), which is scalable to hundreds of independent inputs and suitable for the processing of signals from "Large-N" arrays of many radio antennas. The computational part of the algorithm, the X-engine, is implemented efficiently on Nvidia's Fermi architecture, sustaining up to 79% of the peak single precision floating-point throughput. We compare

performance obtained for hardware- and software-managed caches, observing significantly better performance for the latter. The high performance reported involves use of a multi-level data tiling strategy in memory and use of a pipelined algorithm with simultaneous computation and transfer of data from host to device memory. The speed of code development, flexibility, and low cost of the GPU implementations compared to ASIC and FPGA implementations have the potential to greatly shorten the cycle of correlator development and deployment, for cases where some power consumption penalty can be tolerated.

2012

NVIDIA -

NVIDIA's Next Generation CUDA™ Compute Architecture: Kepler™ GK110

V. K. Veligatla, P. Labropoulos, L. V. E. Koopmans -

Adaptive Beam-forming for Radio Astronomy On GPU

The LOFAR radio telescope consists of tens of thousands of dipole antennas that combine their signals to operate as a single large radio telescope. The truly innovative aspect of this new telescope is that its pointing system is not mechanical. It is steered by combining the electric signals from different elements using advanced beam-forming software. Imaging software is one of the important aspects of processing the high-volume data streams produced by LOFAR, and is one of the best places to use GPUs to achieve processing speed. We were able to achieve up to 30 times performance gain compared to the CPU implementation in novel, computationally intensive techniques such as the Minimum Variance Distortionless Response (MVDR). We have gained 5-6 times speed-up compared to the CPU implementation for standard imaging algorithms.

John W. Romein -

An Efficient Work-Distribution Strategy for Gridding Radio-Telescope Data on GPUs

This paper presents a novel work-distribution strategy for GPUs, that efficiently convolves radio-telescope data onto a grid, one of the most time-consuming processing steps to create a sky image. Unlike existing work-distribution strategies, this strategy keeps the number of device-memory accesses low, without incurring the overhead from sorting or searching within telescope data. Performance measurements show that the strategy is an order of magnitude faster than existing accelerator-based gridders. We compare CUDA and OpenCL performance for multiple platforms. Also, we report very good multi-GPU scaling properties on a system with eight GPUs, and show that our prototype implementation is highly energy efficient. Finally, we describe how a unique property of GPUs, fast texture interpolation, can be used as a potential way to improve image

quality.

Alessio Sclocco, Ana Lucia Varbanescu, Jan David Mol, Rob V. van Nieuwpoort - Radio Astronomy Beam Forming on Many-Core Architectures

Traditional radio telescopes use large steel dishes to observe radio sources. The largest radio telescope in the world, LOFAR, uses tens of thousands of fixed, omnidirectional antennas instead, a novel design that promises ground-breaking research in astronomy. Where traditional telescopes use custom-built hardware, LOFAR uses software to do signal processing in real time. This leads to an instrument that is inherently more flexible. However, the enormous data rates and processing requirements (tens to hundreds of teraflops) make this extremely challenging. The next-generation telescope, the SKA, will require exaflops. Unlike traditional instruments, LOFAR and SKA can observe in hundreds of directions simultaneously, using beam forming. This is useful, for example, to search the sky for pulsars (i.e. rapidly rotating highly magnetized neutron stars). Beam forming is an important technique in signal processing: it is also used in WIFI and 4G cellular networks, radar systems, and health-care microwave imaging instruments. We propose the use of many-core architectures, such as 48- core CPU systems and Graphics Processing Units (GPUs), to accelerate beam forming. We use two different frameworks for GPUs, CUDA and OpenCL, and present results for hardware from different vendors (i.e. AMD and NVIDIA). Additionally, we implement the LOFAR beam former on multi-core CPUs, using OpenMP with SSE vector instructions. We use autotuning to support different architectures and implementation frameworks, achieving both platform and performance portability. Finally, we compare our results with the production implementation, written in assembly and running on an IBM Blue Gene/P supercomputer. We compare both computational and power efficiency, since power usage is one of the fundamental challenges modern radio telescopes face. Compared to the production implementation, our auto-tuned beam former is 45–50 times faster on GPUs, and 2–8 times more power efficient. Our experimental results lead to the conclusion that GPUs are an attractive solution to accelerate beam forming.

V.Vamsi Krishna, Dr. Panos Labropoulos, Prof. Leon V.E. Koopmans - GPU's for Radio Imaging

- *Signals from Sources (e.g. galaxies)*
- *Next Gen Antennas (e.g. LOFAR, SKA, ...)*
- *Image acquired after Processing (RFI elimination, Calibration).*

Mike Clark with Lincoln Greenhill and Paul LaPlante - Accelerating Radio Astronomy Cross-Correlation Beyond 1 Tflops Using Fermi

2013

Harshavardhan Reddy Suda, Pradeep Kumar Gupta -

Powering Real-time Radio Astronomy Signal Processing with GPUs. Design of a GPU based real-time backend for the upgraded GMRT

Nitisha Pirthee -

Digital back end for MITRA prototype

In the first part of the project, USRP1 was used on GNU radio. A log periodic antenna was connected to one channel of the USRP and the expected peaks were observed. The second channel was not operational when tested. An array of sixteen channels was used as front end. A PCI-ADC card already available at MRT was used to do data acquisition. The program for data acquisition was improved. The card was tested several times and observations were carried out. Celestial objects like CAS A and Pictor A were successfully observed.

Harshavardhan Reddy Suda, Pradeep Kumar Gupta -

Powering Real-time Radio Astronomy Signal Processing with GPUs. Design of a GPU based real-time backend for the upgraded GMRT

Ben Barsdell, Mike Clark, Lincoln Greenhill, Jonathon Kocz -

ACCELERATING RADIO ASTRONOMY CROSS-CORRELATION USING THE KEPLER ARCHITECTURE

Kepler GK110 optimisation

2014

Ben Barsdell, Mike Clark, Lincoln Greenhill, Jonathon Kocz -

PETASCALE CROSS-CORRELATION

Amr H. Hassan, Christopher Fluke, David Barnes, Virginia Kilborn -

Astronomical “Big Data” Analysis and Visualization

Alex Bogert, John Holdener, and Nicholas Smith -

Interactive Visualization of Astrophysical Data

yt [1] is an analysis and visualization system for astrophysical volumetric data that is openly developed and freely available. At its core, yt provides a method of describing physical rather than computational objects inside an astrophysical simulation. yt provides methods for selecting regions, applying analysis to regions, visualizing (including volume rendering, projections, slices,

phase plots) and exporting data to external analysis packages.

S. Bhatnagar,P. K. Gupta, M. Clark -
GPU based imager for radio astronomy

Mario Guillaume CECILE -

Enhancement of some computational physics algorithms using Parallel Computing and the Graphical Processing Unit

Scientific computing has become an important method for testing and improving current scientific models and theories. Recent developments in computer architecture have helped to study more complex systems using High Performance Computing (HPC). In this project, Grain Growth simulation and the soft-sphere Discrete Element Method (DEM) are enhanced to be able to consider larger matrix sites in the Grain Growth simulation and a large number of particles in the DEM model. Parallel computing using Message Passing Interface (MPI) is used as well as CUDA for programming on the GPU. For the Grain Growth simulation, effects of foreign particles are investigated while for the DEM model, free-falling of particles in a packed bed is studied. The results presented in this work help to give further understanding about the physics involved behind both the Grain Growth and DEM. This work furthermore demonstrates how the use of parallel processing can help scientists to enhance their code.

NVIDIA -

NVIDIA GeForce GTX 750 Ti Featuring First-Generation Maxwell GPU Technology,
Designed for Extreme Performance per Watt

Chapter 3

Cross-correlation

3.1 The correlator

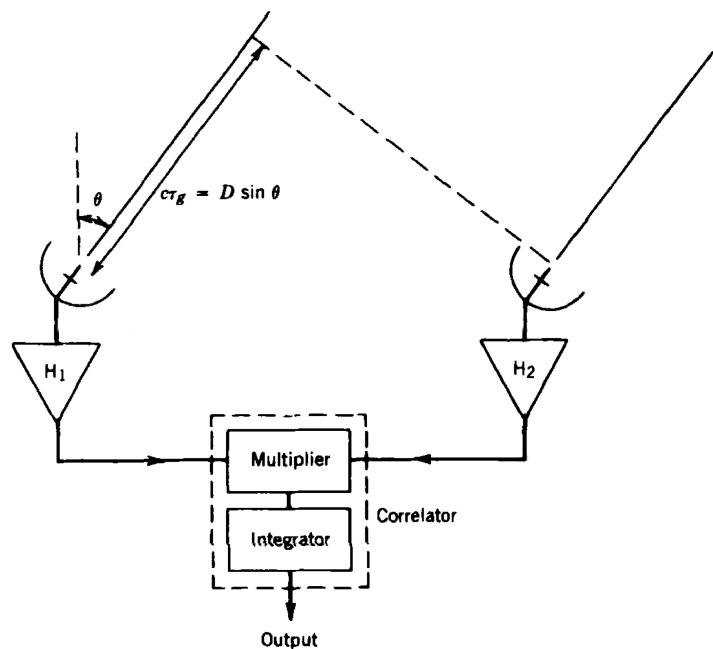


FIGURE 3.1:
A simplistic representation of a correlator [6, Pg. 53, Fig. 2.3]

[From 6, Sec 2.2] One might ask, how do we get the values of visibilities? Remember that in chapter 1, section 1.3.3, last paragraph we said that it is a measure that characterises the relationship between the antenna-pair signal. Basically to obtain this measure, the antenna-pair signals is fed to a device which processes the inputs and outputs a signal which relates the input signals, there exist different types of devices that one might use to obtain that information, i.e., the visibility, in our case we will focus on the use of a correlator.

Basically a correlator does the following things as illustrated in figure 3.1 it multiplies the signals

together and integrates over a certain interval of time called the integration time. Now assume that for a point source each antenna delivers the same signal, i.e. the voltage, $V(t)$, to the correlator and that one lags the other by a time delay, τ , the output of the correlator may be a voltage, a current, or a coded set of logic levels (for a digital system), but in any case it represents a physical quantity with the dimensions of voltage squared. And we can express it as in the following equation 3.1,

$$r(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V(t)V(t - \tau)dt \quad (3.1)$$

This is actually an autocorrelation function. The signal from a natural cosmic source can be considered as a continuous random process that results in a broad spectrum, of which the phases are a random function of frequency. Assume that the time-averaged amplitude of the cosmic signal in any finite band is constant with frequency over the passband of the receiver. The squared amplitude of a frequency spectrum is known as the power density spectrum, or power spectrum. The power spectrum of a signal is the Fourier transform of the autocorrelation function of that signal. This statement is known as the Wiener-Khinchin relation which can be written as the following,

$$|H(\nu)|^2 = \int_{-\infty}^{\infty} r(\tau) e^{-j2\pi\nu\tau} d\tau \quad (3.2)$$

and its pair,

$$r(\tau) = \int_{-\infty}^{\infty} |H(\nu)|^2 e^{j2\pi\nu\tau} d\nu \quad (3.3)$$

3.2 Cross-correlation

[From 6, Sec. 3.3] It is also useful to examine the corresponding relation for the **cross-correlation function** of two different waveforms. The response of a correlator, as used in a radio interferometer, can thus be written as

$$r(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V_1(t)V_2^*(t - \tau)dt \quad (3.4)$$

for which we have the pentagram as a short hand notation, as follows,

$$V_1(t) \star V_2(t) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T V_1(t)V_2^*(t - \tau)dt \quad (3.5)$$

From the convolution theorem one can very easily derive the following relationship,

$$V_1(t) \star V_2(t) \rightleftharpoons \widehat{V}_1(\nu) \cdot \widehat{V}_2^*(\nu) \quad (3.6)$$

Where in all the previous and following equations the superscript asterisk * denotes complex conjugation and, \rightleftharpoons , denotes Fourier transform.

Now let's get back to equation 3.4, τ , is the time by which voltage V_2 is delayed with respect to voltage V_1 . The functions V_1 , and V_2 that represent the signals in equation 3.4 may be complex. The output of a single multiplying device is a real voltage or number though. To obtain the complex cross-correlation, which represents both the amplitude and the phase of the visibility, one can record the fringe oscillations and measure their phase, or use a complex correlator which contains two multiplying circuits.

For the case where the antennas track the source, both the antenna beam center and the center of the source are at the (l, m) origin, the correlator output can thus also be expressed as the following,

$$r(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(l, m) A(l, m) |H(\nu)|^2 e^{j2\pi\nu\tau} dl dm d\nu \quad (3.7)$$

For a wavefront incident from the direction (l, m) , the difference in propagation times through the two antennas to the correlator results from a difference in path lengths of $(ul + vm)$ wavelengths, where an approximation has been made (refer to the statement before and after Eq. 1.16). The corresponding time difference is $\frac{(ul+vm)}{\nu}$. If we take as V_1 , the signal from the antenna for which the path length is the greater (for positive l and m), then from equation 3.7, the correlator output becomes

$$r(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(l, m) A(l, m) |H(\nu)|^2 e^{-j2\pi(lu+mv)} dl dm d\nu \quad (3.8)$$

Assuming that the intensity and the antenna pattern are constant over the bandpass range of the filters, and the width of the source is small compared with the antenna beam. The correlator output then becomes

$$r = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(l, m) A(l, m) dl dm \int_{-\infty}^{\infty} |H(\nu)|^2 e^{-j2\pi(lu+mv)} d\nu \quad (3.9)$$

$$= A_0 V(u, v) \int_{-\infty}^{\infty} |H(\nu)|^2 e^{-j2\pi(lu+mv)} d\nu \quad (3.10)$$

where A_0 is the collecting area of the antennas in the direction of the maximum beam response and $V(u, v)$ the true visibility, the measured visibility, $V'(u, v)$ was introduced in chapter 1 in the equation 1.17. The filter response $H(\nu)$ is a dimensionless (gain) quantity, note that we assume that the antennas are identical and that their filter response are identical such that $H_1(\nu) = H_2(\nu) = H(\nu)$. If the filter response is essentially constant over a bandwidth, A_0 , eq. 3.10 becomes

$$r = A_0 V(u, v) \Delta\nu \quad (3.11)$$

Thus we have here the visibility $V(u, v)$ which has units of $\text{Wm}^2\text{Hz}^{-1}$, A_0 has units of m^2 , and $\Delta\nu$ has units of Hz. This is consistent with r , the output of the correlator, which is proportional to the correlated component of the received power.

Chapter 4

Image Reconstruction

4.1 The Van Cittert-Zernike theorem

As seen in chapter 1 and chapter 3, the interferometer measures the complex visibility, $V'(u, v)$, of a source, which is the Fourier transform of its intensity distribution multiplied by the primary beam response as observed in the equations, 1.17 and 3.9. The true visibility, $V(u, v)$ can be expressed as follows [4, Slide 8]:

$$V(u, v) = |V|e^{-j\phi} = \iint I(l, m)e^{-j2\pi(ul+vm)}dl dm \quad (4.1)$$

The Fourier transform relationship between the true visibility, $V(u, v)$, and the sky intensity distribution, $I(l, m)$, is the Van Cittert-Zernike theorem on which synthesis imaging is based. This means that we can recover $I(l, m)$ from $V(u, v)$:

$$V(u, v) = \iint I(l, m)e^{-2j\pi(ul+vm)}dl dm \quad (4.2)$$

$$I(l, m) = \iint V(u, v)e^{2j\pi(ul+vm)}dudv \quad (4.3)$$

In eq. 4.1, the visibility is expressed as, $|V|e^{-j\phi}$. The phase, ϕ , contains information about the

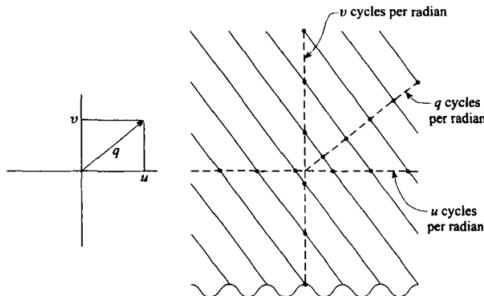


FIGURE 4.1: Fringe Visibility [6, Pg. 64, Fig. 2.7]

location of structure with spatial frequency (u, v) relative to the phase centre ($l = 0, m = 0$), and the amplitude, $|V|$, gives information about how much of the spatial frequency component is present [4, Slide 8] [1, Slide 9].

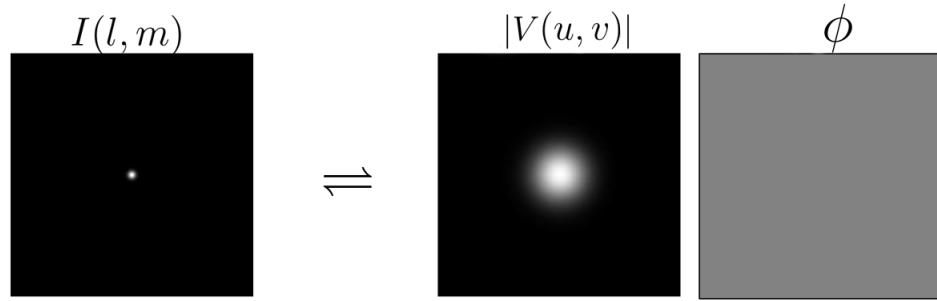


FIGURE 4.2: Amplitude and fringe of the visibility function, source at phase centre [1, Slide 9]

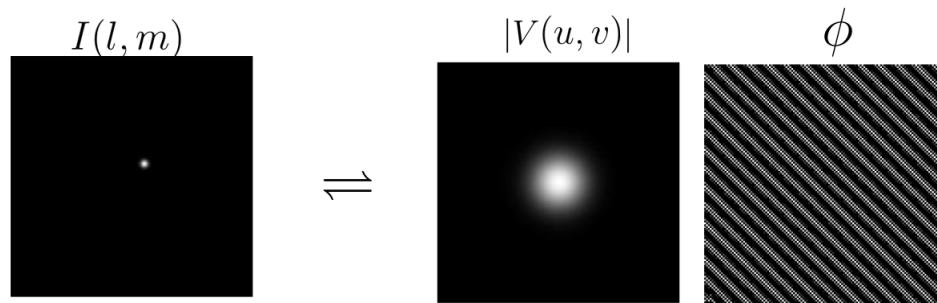


FIGURE 4.3: Amplitude and fringe of the Visibility function source offset from the phase centre in the upper-right direction [1, Slide 9]

4.2 Sampling of the visibility

We already know about the visibility from the previous sections, and chapter 3 which showed how it can be obtained for a pair of antennas. Now consider the following, let us denote the baseline vector components for a particular interferometer antenna-pair, that we index (i, k) , as (u_{ik}, v_{ik}) . For a single measurement of cross-correlation of the voltages, $V_i(t)$, and, $V_k(t)$, from the interferometer, e.g. $V_i \star V_k$, we get a value which relates to the visibility value, $V(u_{ik}, v_{ik})$, however though a fairly important property of the visibility function, $V(u, v)$, is that it is hermitian [6, Pg. 138, Sec. 5.4], so for the example we mentioned, we can write the following:

$$V(u_{ik}, v_{ik}) = V^*(-u_{ik}, -v_{ik}) \quad (4.4)$$

Thus for one measurement with a particular pair of antennas we can obtain data about $V(u, v)$ at two positions on the $u - v$ plane as one can see from equation 4.4, its like the interferometer acts as a filter [6, Pg. 133, Sec. 5.3] that responds to spatial frequencies (u_{ik}, v_{ik}) and $(-u_{ik}, -v_{ik})$.

Ultimately what we want is to obtain an appropriate coverage of the u - v plane, so we need to have different baselines to be able to accomplish this, usually this is done by having an array of antennas, or by having adjustable baselines or/and also using earth rotation synthesis which we'll discuss later, and the choice of configuration of the antennas of a synthesis array is based on optimising, in some manner, the sampling of the visibility function in (u, v) space [6, Pg. 126, Sec. 5.2].

4.2.1 Sampling Theorem

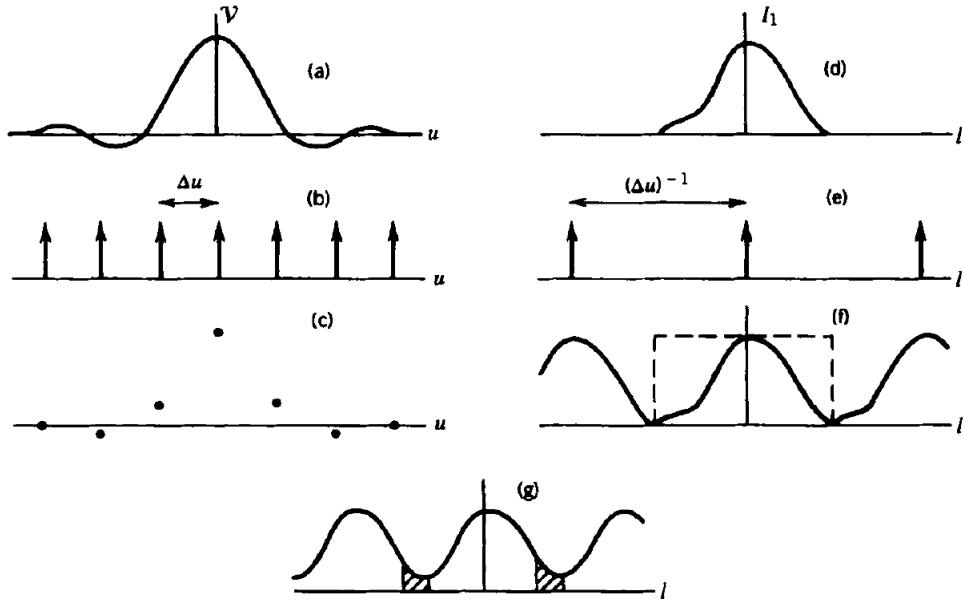


FIGURE 4.4: Sampling a 1-D visibility [6, Pg. 126, Fig. 5.2]

[From 6, Sec. 5.2] Now to appreciate the effect of sampling and its requirement let's focus on the following example, in which one wants to measure the one-dimensional intensity distribution of a source, $I(l, 0)$, portrayed in Fig. 4.4(d). It is necessary to measure the complex visibility, V , in the corresponding direction on the ground at a series of values of the projected antenna spacing. For example, to measure an east-west profile, a possible method is to make observations near meridian transit of the source using an east-west baseline, and to vary the length of the baseline from day to day [6, Pg. 126], or have an array of antennas with baselines which have multiples of a unit spacing, Δu . So we can represent the sampling as the multiplication of the visibility function with the impulse train function (familiar to the students), $\text{III}_{\Delta u}(u)$, defined as follows,

$$\text{III}_{\Delta u}(u) = \sum_{i=-\infty}^{\infty} \delta(u - i\Delta u) \quad \text{Wikipedia [7]} \quad (4.5)$$

Thus the sampled, V , corresponds to the following,

$$S(u) \cdot V(u, 0) = \text{III}_{\Delta u}(u) \cdot V(u, 0) = \sum_{i=-\infty}^{\infty} \delta(u - i\Delta u) \cdot V(u, 0) \quad (4.6)$$

where, S , represents the general sampling function, now the Fourier transform of the impulse function is the following [6, Pg. 127, Eq. 5.6],

$$\text{III}_{\Delta u}(u) \rightleftharpoons \text{III}_{\Delta u^{-1}}(l) = \sum_{i=-\infty}^{\infty} \delta(l - \frac{i}{\Delta u}) \quad (4.7)$$

as per the usual property of scaling for Fourier transforms. And from the convolution theorem.

$$\text{III}_{\Delta u}(u) \cdot V(u, 0) \rightleftharpoons \text{III}_{\Delta u^{-1}}(l) * I(l, 0) \quad (4.8)$$

[From 6, Pg. 127, Sec. 5.2] The result is the replication of, $I(l, 0)$, at intervals Δu^{-1} as observed in Fig. 4.4(f). If, $I(l, 0)$, represents a source of finite dimensions, the replications of $I(l, 0)$ will not overlap as long as $I(l, 0)$ is nonzero only within a range of l that is no greater than Δu^{-1} . An example of overlapping replications is shown in Fig. 4.4(g). The loss of information resulting from such overlapping is commonly referred to as aliasing, because the components of the function within the overlapping region lose their identity with respect to which end of the replicated function they properly belong. Avoidance of aliasing requires that the sampling interval Δu be no greater than the reciprocal of the interval in l within which $I(l, 0)$ is nonzero. To be precise, we should consider the width of the source as broadened by the finite resolution of the observations, rather than the true width of the source, but this is usually only a minor effect. The requirement for the restoration of a function from a set of samples, for example, deriving the function in Fig. 4.4(a) from the samples in Fig. 4.4(c) is easily understood by considering the Fourier transforms in Fig. 4.4(d) and (f). Interpolation in the u domain corresponds to removing the replications in the l domain, which can be achieved by multiplication of the function in Fig. 4.4(f) by the rectangular function, $\Pi_{\Delta u^{-1}}(l)$ indicated by the broken line. In terms of the heavyside or unit step function, $H(l)$, (more familiar to the student) the rectangular function can be expressed as the following [5],

$$\Pi_{\Delta u^{-1}}(l) = H(l + \frac{\Delta u^{-1}}{2}) - H(l - \frac{\Delta u^{-1}}{2}) \quad (4.9)$$

In the u domain this multiplication corresponds to the convolution of the sampled values with the Fourier transform of the rectangular function,

$$\Pi_{\Delta u^{-1}}(l) \rightleftharpoons \Delta u^{-1} \cdot \text{sinc}(u/\Delta u) \quad (4.10)$$

where, $\text{sinc}(x)$, is the normalised sinc function,

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (4.11)$$

$$\Pi_{\Delta u^{-1}}(l) \Leftrightarrow \frac{\sin(\pi u/\Delta u)}{\pi u} \quad (4.12)$$

If aliasing is avoided, convolution with $\frac{\sin(\pi u/\Delta u)}{\pi u}$ provides exact interpolation of the original function from the samples. Thus we can state, as a sampling theorem for the visibility, that if the intensity distribution is nonzero only within an interval of width, l_w , $I(l, 0)$ is fully specified by sampling the visibility function at points spaced $\Delta u = l_w^{-1}$ in u . In two dimensions, it is simply necessary to apply the theorem separately to the source in the l and m directions [6, Pg. 127].

[From 6, Pg. 128, Sec. 5.2] The discrete form of the Fourier transform is very widely used in

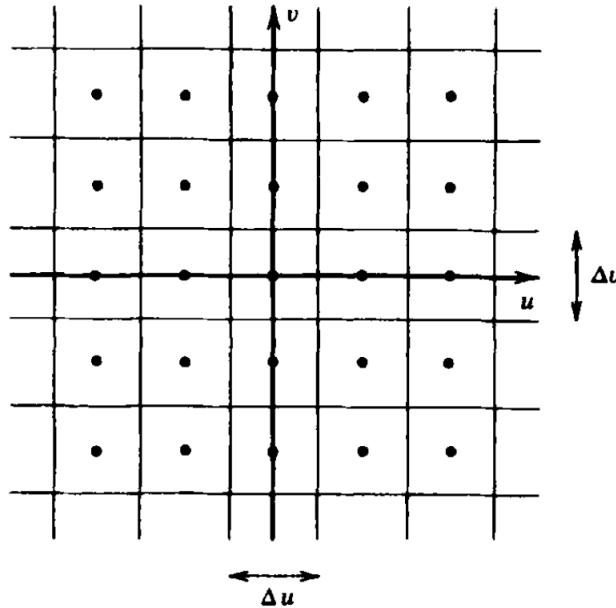


FIGURE 4.5: Sampling a 2-D visibility [6, Pg. 129, Fig. 5.3]

synthesis mapping because of the computational advantages of the fast Fourier transform (FFT) algorithm. With the discrete transform the functions $V(u, v)$ and $I(l, m)$ are expressed as rectangular matrices of sampled values at uniform increments in the two variables involved. The rectangular grid points on which the intensity is obtained provide a convenient form for further data processing. The two-dimensional form of the discrete transform for the Fourier pair $V(u, v)$ and $I(l, m)$ is defined by,

$$V(p\Delta u, q\Delta v) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{k=0}^{N-1} I(i\Delta l, k\Delta m) e^{-j2\pi ip/M} e^{-j2\pi qa/N} \quad (4.13)$$

and the inverse is

$$I(i\Delta l, k\Delta m) = \frac{1}{MN} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} V(p\Delta u, q\Delta v) e^{j2\pi ip/M} e^{j2\pi kq/N} \quad (4.14)$$

The functions are periodic with periods of M samples in the i and p dimensions and N samples in the k and q dimensions. Evaluation of Eqs. 4.13 or 4.14 by direct computation requires approximately $(MN)^2$ complex multiplications. In contrast, if M and N are powers of 2 the FFT algorithm requires only $\frac{1}{2}MN \log_2(MN)$ complex multiplications. The dimensions of the (u, v) plane that contain these data are $M\Delta u$ by $N\Delta v$. In the (l, m) plane the points are spaced Δl in l and Δm in m , and the map dimensions are $M\Delta l$ by $N\Delta m$. The dimensions in the two domains are related by

$$\begin{aligned} \Delta u &= (M\Delta l)^{-1}, & \Delta v &= (N\Delta m)^{-1} \\ \Delta l &= (M\Delta u)^{-1}, & \Delta m &= (N\Delta v)^{-1} \end{aligned} \quad (4.15)$$

The spacing between points in one domain is the reciprocal of the total dimension in the other domain. Thus, if the size of the array in the intensity domain is chosen to be large enough that the intensity function is nonzero only within the area $M\Delta l \times N\Delta m$, then the spacings Δu and Δv in Eq. 4.15 satisfy the sampling theorem.

4.2.2 Array configuration consideration

An array of antennas can be interconnected to operate as a correlator array. For n_a antennas there are $n_a(n_a - 1)$, voltage cross-product terms of form $V_i V_k$ involving different antennas i and k , and, n_a voltage self product terms of form V_i^2 [6, Pg. 130, Sec. 5.3].

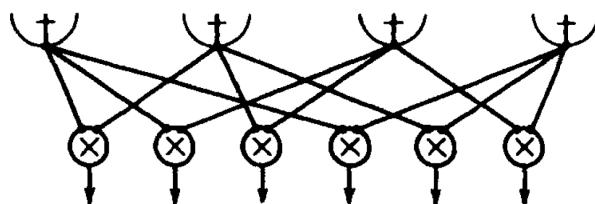


FIGURE 4.6: A correlator array [6, Pg. 130, Fig. 5.4(b)]

4.2.2.1 Spatial Sensitivity and the Spatial Transfer function

[From 6, Sec. 5.3] We can now consider the sensitivity of an antenna array to the spatial frequencies on the sky. The angular response pattern see figure 1.6 of an antenna is the same in reception or transmission, so let's consider the antenna in transmission here, then the power applied to the terminals produces a field at the antenna aperture. A function $W(u, v)$ is equal to the autocorrelation function of $\mathcal{E}(x_\lambda, y_\lambda)$, the distribution of the electric field across the aperture, where x_λ , and y_λ are coordinates in the aperture plane of the antenna and are measured in wavelength, (this is similar to what we did for the Wiener-Khinchin relation in section 3.1 and if we apply the form of the cross-correlation eqs. in section 3.2 for an autocorrelation). Thus

$$W(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{E}(x_\lambda, y_\lambda) \star \star \mathcal{E}^*(x_\lambda - u, y_\lambda - v) dx_\lambda dy_\lambda \quad (4.16)$$

$W(u, v)$ is a very useful or important function, it is proportional to the number of ways, suitably weighted by the field intensity, in which a specific spacing vector (u, v) can be found within the antenna aperture. In reception it is a measure of the sensitivity of the antenna to different spatial frequencies, this is why $W(u, v)$ is also referred to as the transfer function. If we consider the response of the array to a **point source** as for a point source the visibility is constant over the whole (u, v) plane, the measured spatial frequencies are proportional to $W(u, v)$. Thus the point source response $\mathcal{A}(l, m)$ is the Fourier transform of $W(u, v)$. Where $\mathcal{A}(l, m) = A(-l, -m)$ but as the function is usually symmetrical it is not of great practical importance here.

Figure 4.7 shows some commonly used configurations of antenna arrays see fig. 4.7 (a),(c),(e),

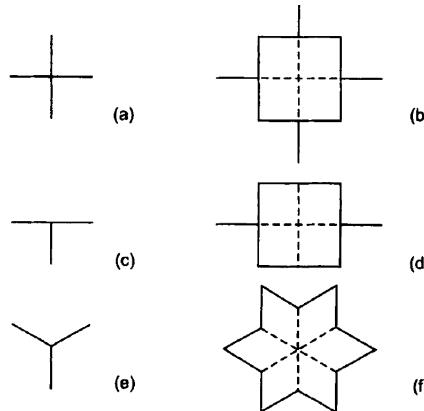


FIGURE 4.7: Antenna apertures configuration and autocorrelation nonzero boundaries [6, Pg. 136, Fig. 5.7]

and the boundaries (b),(d),(f) of their autocorrelation functions. The autocorrelation function indicate the instantaneous spatial sensitivity for a continuous aperture in the form of the corresponding figure. Ridges of high autocorrelation are emphasised by the broken lines. These occur for displacements at which the arms of figures such as those in Figure 4.7 are aligned.

A cross and its autocorrelation function are shown in Fig. 4.7(a) and (b). It is assumed that the width of the arms is finite but small compared with the length of the arms. The spatial sensitivity is represented by the square in 4.7(b). Notice the narrow extensions on the centres of the sides of the square represent parts of the autocorrelation functions of the individual arms, which are not formed in the cross-correlation of the arms. However, they are formed if the arms consist of lines of individual antennas for which the cross-correlation is formed for pairs on the **same arm** as well as those on crossed arms. The case for a T-shaped array is similar and is shown in Fig. 4.7(c) and (d). Again, if only the cross-correlation between the east-west arm and the half-length, north-south arm is formed, then the spatial frequency coverage is represented by the square component of the autocorrelation. The equivalence between the spatial transfer function of such a cross and a T can be understood by noting that for any pair of points in the aperture of a cross, for example, one on the east arm and one on the north arm, there is a corresponding pair on the west and south arms for which the spacing vector is identical. Thus any one of the four half-length arms can be removed without reducing the (u, v) coverage of the spatial transfer function.

An alike example of a non-tracking T configuration is, the Mauritius Radio Telescope, near Bras d'eau, Mauritius, which was a T-shaped array of helix antennas operating at 150 MHz. The east-west arm is 2 km long. The south arm is 880 m long and was synthesised by moving a group of antennas on trolleys in steps, with continuous coverage, to simulate a larger aperture. The spatial frequency coverage is the same as would be obtained in a single observation with an antenna of aperture equal to that simulated by the movement of the antennas, although the magnitude of the spatial sensitivity is not exactly the same [6, Sec. 5.6 Pg. 155, Sec. 5.3 Pg. 137].

Now consider the case of a tracking array, as the source moves in hour angle, the changing (u, v) coverage is represented by a band centred on the spacing locus of the two antennas. Since $V(-u, -v) = V^*(u, v)$, any pair of antennas measures visibility along two arcs symmetric about the (u, v) origin, both of which are included in the spatial transfer function. Because the antennas track the source, the antenna beams remain centred on the same point in the source under investigation, and the array measures the product of the source intensity distribution and the antenna pattern.

To accommodate the effects that result when the antennas track the source position, the normalised antenna pattern is treated as a modification to the intensity distribution, the intensity distribution then becomes $A_N(l, m)I(l, m)$. The spatial transfer function $W(u, v)$ for a pair of tracking antennas is represented at any instant by a pair of two-dimensional delta functions $\delta(u, v)$ and $\delta^*(-u, -v)$. For an array of antennas the resulting spatial transfer function is represented by a series of delta functions weighted in proportion to the magnitude of the instrumental response. As the earth rotates, these delta functions generate the ensemble of elliptical spacing loci. The loci represent the spatial transfer function of a tracking array.

Consider observation of a source $I(l, m)$, for which the visibility function is $V(u, v)$, with normalised antenna patterns $A_N(l, m)$. Then if $W(u, v)$ is the spatial transfer function, the measured visibility is

$$[V(u, v) * * \overline{A_N}(u, v)]W(u, v) \quad (4.17)$$

where the double asterisk indicates two-dimensional convolution and the bar denotes the Fourier transform. The Fourier transform of 4.17 gives the measured intensity:

$$[I(l, m)A_N(l, m)] * * \overline{W}(l, m) \quad (4.18)$$

If we observe a **point source** at the (l, m) origin, where $A_N = 1$, expression 4.18 becomes the point-source response $b_0(l, m)$. We then obtain

$$b_0(l, m) = [\delta(l, m)A_N(l, m)] * * \overline{W}(l, m) = \overline{W}(l, m) \quad (4.19)$$

where $\delta(l, m)$ represents the point source. Here again, the point-source response is the Fourier transform of the spatial transfer function. In the tracking case the spatial frequencies that contribute to the measurement are represented by $W(u, v) * * \overline{A_N}(u, v)$. We also note that $\overline{A_N}(u, v)$ is twice as wide as the corresponding antenna aperture in the (x, y) domain.

As a first step in considering the layout of the antennas it is useful to consider the desired spatial (u, v) coverage. For any specific observation, the optimum (u, v) coverage clearly depends on the expected intensity distribution of the source under study, since one would prefer to concentrate the capacity of the instrument in (u, v) regions where the visibility is nonzero. However, most large arrays are used for a wide range of astronomical objects, so some compromise approach is required. Since, in general, astronomical objects are aligned at random in the sky, there is no preferred direction for the highest resolution. Thus it is logical to aim for visibility measurements that extend over a circular area centred on the (u, v) origin. As described in Section 5.2, the visibility data are usually interpolated onto a rectangular grid for convenience in Fourier transformation, and if approximately equal numbers of measurements are used for each grid point, they can be given equal weights in the transformation. Uneven weighting results in loss of sensitivity, since some values then contain a larger component of noise than others. From this viewpoint one would like the natural weighting (i.e., the weighting of the measurements that results from the array configuration without further adjustment) to be as uniform as possible within the circular area.

Consider a circular (u, v) area of diameter a_λ wavelengths in which there are no holes in the data; that is, the visibility data interpolated onto a rectangular grid for Fourier transformation has no missing values. Then for uniform weighting, the synthesised beam, which is obtained from the Fourier transform of the gridded transfer function has the form of the following bessel function, $J_1(\pi a_\lambda \theta)/\pi a_\lambda \theta$. Let us refer to the (u, v) area described above as the complete (u, v) coverage and the resulting beam as the complete response. Now if some data are missing, the actual (u, v)

coverage is equal to the complete coverage minus the (u, v) hole distribution. By the additive property of Fourier transforms, the corresponding synthesised beam is equal to the complete response minus the Fourier transform of the hole distribution. The holes add an unwanted component to the complete response, in effect adding sidelobes to the synthesized beam.

To wrap up this section, the following illustrations in Fig. 4.8 will help us appreciate how the (u, v) map coverage results in more information about the Intensity distribution

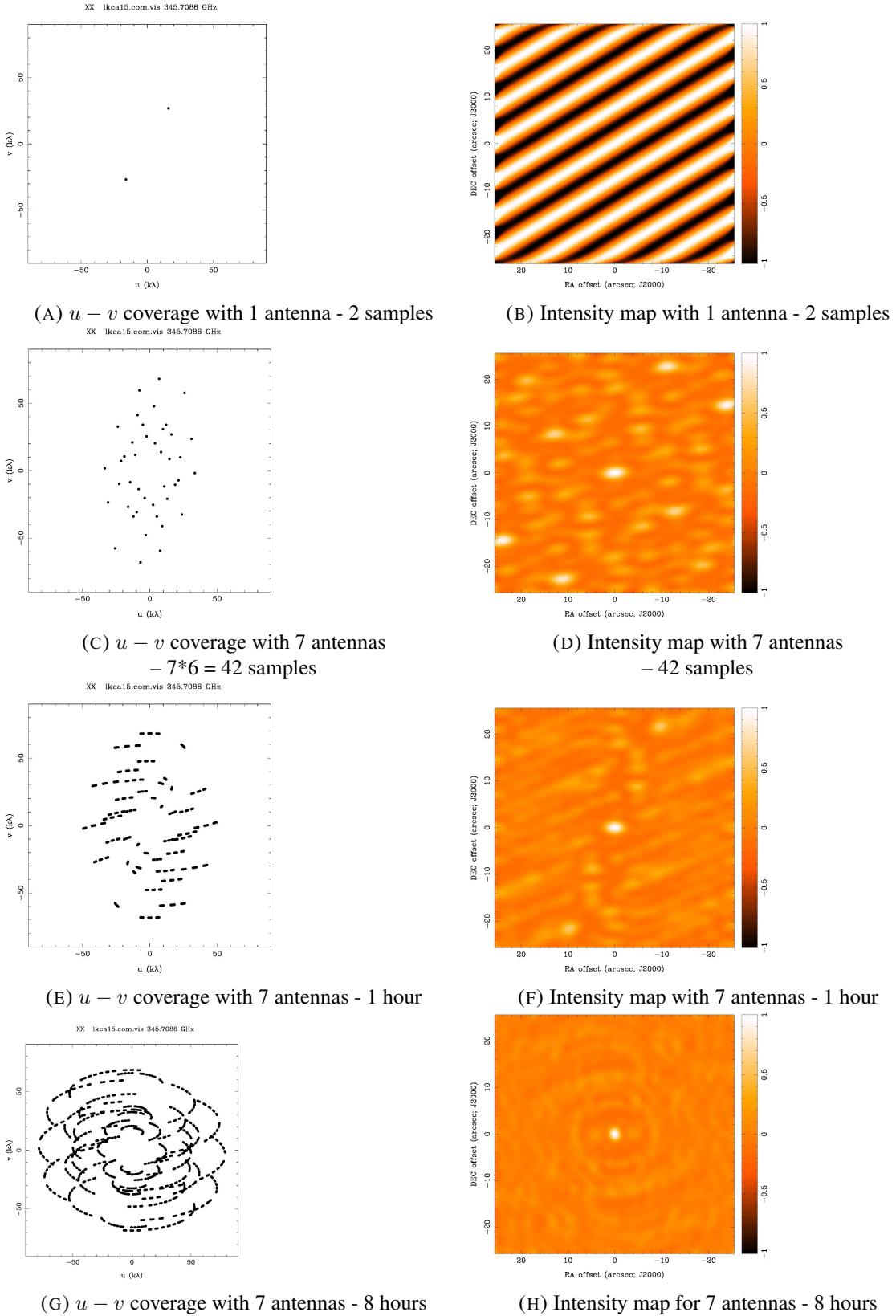


FIGURE 4.8: Visibility map coverage and derived Intensity map [1, Slide 28,33,36,38]

4.3 Data calibration

[From 6, Sec. 10.1 Pgs. 383-385] The aim of calibration is to remove, as much as possible, the effects of instruments and atmospheric factors in instruments. Such factors depend largely on the individual antennas or antenna pairs and their associated electronics, so a correction must be applied to the visibility data before they are combined into an image. Editing the visibility data to delete any that show evidence of radio interference or equipment malfunction is usually performed before the proper calibration. This mainly implies examining samples of data for unexpected levels or phase variations. Data taken on calibration sources are particularly useful here since the response to such a source is predictable and is expected to vary only slowly and smoothly with time.

First instrumental factors that are stable over periods of weeks or more are considered, there are also effects that vary during an observation and principally involves correction of the complex gain of the antenna pairs, these can be divided into two categories, the ones which can be predicted or measured and those which must be determined by observing a calibration source during the observation period.

4.3.1 Calibration using calibration sources

[From 6, Sec. 10.1 Pgs. 385-387] The steps in calibration involve parameters that may vary on timescales of minutes or hours and require the observation of one or more calibration sources. The source that is the subject of the astronomical investigation is referred as the target source to distinguish it from the calibration source, or calibrator. From Eq. 1.17 we can write the expression for the interferometer response as follows:

$$[V(u, v)]_{uncal} = G_{ik}(t) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{A_N(l, m) I(l, m)}{\sqrt{1 - l^2 - m^2}} e^{-j2\pi(ul + vm)} dl dm \quad (4.20)$$

Where, $[V(u, v)]_{uncal}$, is the uncalibrated visibility and, $I(l, m)$ is the source intensity. The complex gain factor $G_{ik}(t)$ is a function of the antenna pair (i, k) and, as a result of unwanted effects, may vary over time. A_N is the antenna aperture normalised to unity for the direction of the main beam. It can be removed from the source image as a final step in image processing and this is discussed in chapter 5. The factor $A_N(l, m)/\sqrt{1 - l^2 - m^2}$ in the intensity-visibility relationship is close to unity, and thus here on we shall omit it, except in the case of wide-field mapping that we won't cover in this report. To calibrate $G_{ik}(t)$, an unresolved calibrator can be observed for which the measured response is

$$V_c(u, v) = G_{ik}(t) S_c \quad (4.21)$$

where the subscript, c , is to denote the calibrator and S_c is the flux density of the calibrator. In calibrating it is best to consider the amplitude and phase separately, since the errors in these two quantities generally arise through different mechanisms. To calibrate the visibility of the target source we can thus write

$$V(u, v) = \frac{[V(u, v)]_{uncal}}{G_{ik}(t)} = [V(u, v)]_{uncal} \left[\frac{S_c}{V_c} \right] \quad (4.22)$$

To observe the calibration source it is placed at the phase centre of its field. Then assuming that the calibrator is unresolved, the phase is a direct measure of the instrumental phase. Thus phase calibration for the target source simply requires subtracting the calibrator phase from the observed phase. The visibility amplitude can be calibrated using the moduli of the visibility terms in Eq. 4.22. The response to the calibrator should be corrected for the calculable and/or directly monitored effects before the gain calibration is performed. Where there are separate receiving channels for two opposite polarizations at each antenna, the calibration must be performed separately for each one. For measurements of source polarization further calibration procedures are necessary which will not be discussed here. The complex gain factor $G_{ik}(t)$ is of the antenna pair (i, k) is indicated as follows,

$$G_{ik}(t) = g_i g_k^* \quad (4.23)$$

so the measured gains for antenna pairs can be used to determine gain factors for the individual antennas. Using the antenna gain factors rather than the baseline gain factors reduces the calibration data to be stored, and helps in monitoring the performance of individual antennas. Also, with this technique, some of the spacings can be omitted from the calibration observation so long as each of the antennas is included. In practice, gain tables including both amplitude and phase are generated for the antennas as a function of time, and the values are interpolated to the times at which data from the target source were taken. The interpolation should be done separately for the amplitude and phase, not for the real and imaginary parts of the gain, otherwise the phase errors can degrade the amplitude, and vice versa. The desirable characteristics of a calibration source are the following.

- Flux density – The calibrator should be strong, so that a good signal-to-noise ratio is obtained in a short time, to reduce the (u, v) coverage lost from the target source.
- Angular width – The calibrator should, if possible, be unresolved so that precise details of its visibility are not required.
- Position – The position of the calibrator should be close to that of the target source. Effects in the atmosphere or antennas that cause the gain to vary with pointing angle are then more effectively removed, and time lost in driving the antennas between the target source and calibrator positions is kept small.

It is not always possible to find a calibrator that satisfies all of the above requirements. In such cases it may be necessary to find a source that is largely unresolved and close to the target source, and then calibrate it against one of the more commonly used flux density references.

4.4 Image Reconstruction

[From 6, Sec. 10.1 Pgs. 387-394] The most straightforward method of obtaining an intensity distribution from measured visibility data is by direct Fourier transformation. The measured visibility $V_{meas}(u, v)$ can be written

$$V_{meas}(u, v) = W(u, v)w(u, v)V(u, v), \quad (4.24)$$

where $W(u, v)$ is the transfer function as introduced in the previous section 4.2.2.1, and $w(u, v)$ represents any applied weighting. The Fourier transform of Eq. 4.24 is the measured intensity distribution, which is

$$I_{meas}(l, m) = I(l, m) * * b_0(l, m), \quad (4.25)$$

Where the double asterisk indicates convolution in 2-D and b_0 is the synthesised beam, which is the Fourier transform of the weighted transfer function:

$$b_0(l, m) \rightleftharpoons W(u, v)w(u, v) \quad (4.26)$$

where, \rightleftharpoons , indicates the Fourier transform relationship. Effects such as non-coplanar baselines and others of minimal importance here and are not included. The visibility is measured at an ensemble of n_d pairs of points symmetric about the (u, v) origin, and the direct Fourier transform of these data is represented by

$$\sum_{i=1}^{n_d} w_i [V_{meas}(u_i, v_i) e^{j2\pi(u_i l + v_i m)} V_{meas}(u_i, v_i) e^{-j2\pi(-u_i l - v_i m)}] \quad (4.27)$$

The weighting factor w_i is introduced to control the form of the synthesized beam, $b_0(l, m)$. Since the visibility at $(-u_i, -v_i)$ is the complex conjugate of the visibility at (u_i, v_i) , the derived intensity is real (here we are considering the case where the antennas are identically polarized.) In the Fourier transformation of the visibility, the intensity is usually computed at points in a rectangular grid with uniform increments in l and m , since this is a very convenient form for subsequent processing.

4.4.1 Weighting of the visibility

To obtain the best signal-to-noise ratio in the summation of measurements that contain Gaussian noise, the individual data values should be weighted inversely as their variance, this is known as natural weighting, for most arrays though natural weighting results in poor beam shape because the shorter spacings are overemphasised. Thus the usual approach is to include in the weighting a factor that is inversely related to the area density of the data in the (u, v) plane. By this mean we obtain data that has a uniform density. However with uniform weighting the strong, near-in sidelobes (close to the main beam) obscure low-level detail and thereby reduce the range of intensity levels that can be reliably measured. The near-in sidelobes of the functions can be reduced at the expense of some increase in the width of the synthesised beam by introducing a Gaussian or similar taper into the weighting function, and thus our resulting weighting function consists of the following

$$w(u, v) = w_u(u, v) \cdot w_t(u, v) \quad (4.28)$$

$w_u(u, v)$ denotes weighting required to obtain uniform effective density, and $w_t(u, v)$ denotes the tapering function. Thus the synthesised beam is the Fourier transform of $\bar{W}(u, v)w_u(u, v)w_t(u, v)$:

$$b_0(l, m) = \overline{\bar{W}}(l, m) * * \overline{w}_u(l, m) * * \overline{w}_t(l, m) \quad (4.29)$$

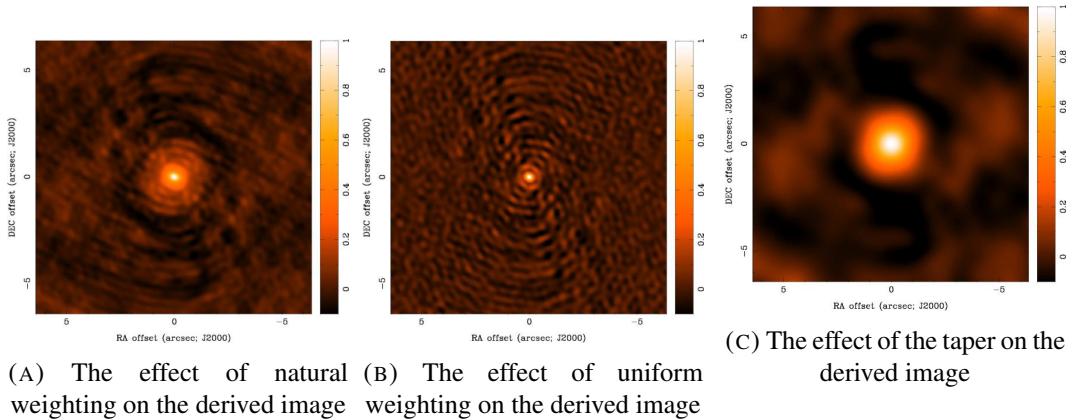


FIGURE 4.9: The effect of weighting visibilities [1, Slide 43,44,46]

where the bar denotes Fourier transform. One thing to note is that the Gaussian taper, $w_t(u, v)$ reduces the sidelobes outside of the main beam at the expense of widening the beam.

4.4.2 Mapping by Discrete Fourier transform

As introduced in the section on sampling, section 4.2.1, the speed of the fast algorithm for the discrete Fourier transform (FFT) is a major advantage in computing large maps. However, the use of the FFT introduces two complications in addition to those discussed for the direct transform: (1)

the necessity to evaluate the visibility at points on a rectangular grid and (2) the resulting possibility of aliasing of parts of the image from outside the synthesised field. So again similarly one has to evaluate the visibility at the grid points, (to denote gridded maps we will use the superscript, G , hereafter,) the output of such a process can be expressed as follows:

$$V^G(u, v) = w(u, v) \text{III}_{(\Delta u, \Delta v)}(u, v) \{C(u, v) * * [W(u, v) \cdot V(u, v)]\} \quad (4.30)$$

Here the visibility $V(u, v)$, measured at the points denoted by the transfer function, $W(u, v)$, is convolved with a function $C(u, v)$, to produce a continuous visibility distribution. This is then resampled at points in a rectangular grid with incremental spacings Δu and Δv . This process is often referred to as **convolutional gridding**. The resampling is here represented by the two-dimensional impulse train function defined by,

$$\text{III}_{(\Delta u, \Delta v)}(u, v) = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} \delta(u - i\Delta u, v - k\Delta v), \quad (4.31)$$

where δ is the two-dimensional delta function. The weighting, $w(u, v)$ to optimize the beam is applied to the resampled data. In practice though, the convolution is evaluated only at the grid points. The Fourier transform of Eq. 4.30 represents the measured intensity:

$$I_{meas}^G(l, m) = \text{III}_{(\Delta u^{-1}, \Delta v^{-1})}(l, m) * * \bar{w}(l, m) * * \{\bar{C}(l, m)[\bar{W}(l, m) * * I(l, m)]\} \quad (4.32)$$

As a result of the Fourier transformation, the intensity function $I(l, m)$ is convolved with the Fourier transform of the transfer function, multiplied by $\bar{C}(l, m)$ which is the Fourier transform of the convolving function, and then convolved with the Fourier transforms of the weighting and resampling functions. This last convolution, i.e. with $\text{III}_{(\Delta u^{-1}, \Delta v^{-1})}(l, m)$, causes the whole map to be replicated at intervals, Δu^{-1} in l and Δv^{-1} in m . These intervals are equal to the dimensions of the map in the (l, m) plane; that is, $\Delta u^{-1} = M\Delta l$ and $\Delta v^{-1} = N\Delta m$, for an $M \times N$ point array. The function $\bar{C}(l, m)$ takes the form of a taper applied to the map, and if this function does not vary greatly on the scale of the width of $\bar{w}(l, m)$, which is usually the case for large maps, then $\bar{w}(l, m)$ in Eq. 4.32 can be convolved directly with $\bar{W}(l, m) * * I(l, m)$, and Eq. 4.32 becomes

$$I_{meas}^G(l, m) \simeq \text{III}_{(\Delta u^{-1}, \Delta v^{-1})}(l, m) * * \{\bar{C}(l, m)[I(l, m) * * b_0(l, m)]\} \quad (4.33)$$

where the synthesised beam $b_0(l, m)$ comes from Eq. 4.29. However, the problem of aliasing remains, and the most effective way to deal with this is to convolve the data in the (u, v) plane with the Fourier transform of a function that, in the (l, m) plane, varies very little over the map and then falls off rapidly at the map edges. We therefore look for a convolving function $C(u, v)$ for which the Fourier transform $I(l, m)$ has these properties, for example one could use of a Gaussian-Sinc function as the convolving function (Thompson et al. [see 6, Sec. 10.2 Pgs. 394-399]).

Chapter 5

Image Processing

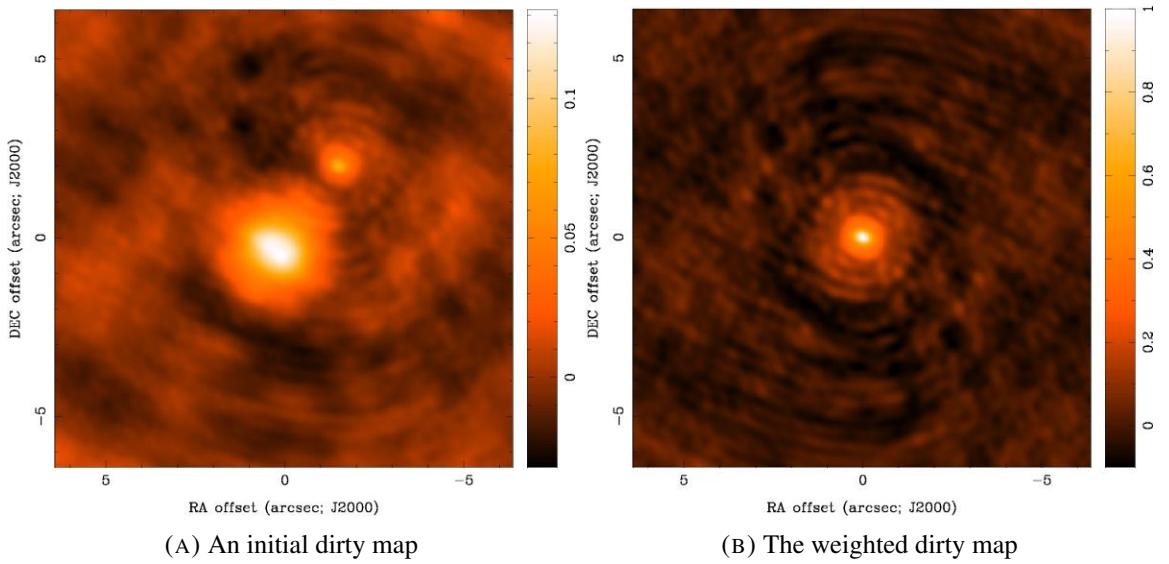


FIGURE 5.1: The end result of image reconstruction [1, Slide 51]

5.1 Introduction - Image Processing

Images obtained from the visibility, $V(u, v)$, corresponding to the antenna responses to the signals from the sky are somehow defective and known as “dirty images” see Fig. 5.1. In order to obtain the highest dynamic range in radio images, both good spatial frequency coverage and effective image processing are required (Thompson et al. [6, Pg. 426]).

Let’s first relate to the principal deficiencies limiting the accuracy of the synthesised images. These are the limited distribution of spatial frequencies in the (u, v) plane and errors in the measurements themselves. One way to improve on the limited spatial frequency coverage is by deconvolution

processes that allow the unmeasured visibility to take nonzero values within some general constraints in the image. While the errors in the measurements themselves can be refined by calibrating the measured visibilities $V'(u, v)$ in order to approximate the true visibilities $V(u, v)$ as closely as possible (Thompson et al. [6, Pg. 426]) which was discussed in chapter 4.

$$I_0(l, m) = I(l, m) * * b_0(l, m) \quad (5.1)$$

$I_0(l, m)$: measured intensity distribution

$I(l, m)$: true intensity

$b_0(l, m)$: synthesized beam

The measured intensity distribution, $I_0(l, m)$, obtained in synthesis mapping can be regarded as the true intensity, $I(l, m)$, convolved with the synthesized beam, $b_0(l, m)$, illustrated there. The double asterisk indicates two-dimensional convolution.

Knowing the measured intensity distribution and the synthesized beam we might want to solve for the true intensity by an analytical procedure for deconvolving two functions that is achieved by taking the Fourier Transform of the convolution, which is equal to the product of the Fourier transforms of the components, dividing out the Fourier transform of the known function, and transforming back [6, Pg. 426].

$$I_0(l, m) \rightleftharpoons [V(u, v)[W(u, v)w_u(u, v)w_t(u, v)] \quad (5.2)$$

$V(u, v)$: visibility function

$W(u, v)$: spatial transfer function

$w_u(u, v)$: weighting required to obtain effective uniform density of data in the (u, v) plane.

$w_t(u, v)$: applied taper

However, the transfer function contains areas where it is zero, so we cannot just divide it out to obtain the visibility function. The holes, that are values at which visibilities are not measured, present a fundamental problem, and any procedure aiming at improving the true intensity other than weighting of the visibility must involve placing nonzero visibility values in the unmeasured (u, v) areas. Therefore as pointed out by Bracewell and Roberts in 1954 there can be an infinite number of solutions to the above convolution, since one can add any arbitrary visibility values in the unsampled areas of the (u, v) plane. The Fourier transform of these added values constitutes an invisible distribution that cannot be detected by any instrument with corresponding zero areas in the transfer function [6, Pg. 427].

It may thus be argued that in interpreting observations from any radio telescope, one should maintain only zeros in the unmeasured regions of spectral sensitivity, to avoid arbitrarily generating information. Alas, zeros are themselves arbitrary values some of which are certainly wrong. What we want is a procedure that allows the visibility at the unmeasured points to take values consistent with the most reasonable or likely intensity distribution, while minimizing the addition of arbitrary detail [6, Pg. 427].

Characteristics such as positivity of intensity and confinement of the angular structure of a source are expected and they can be introduced into the imaging process. Instrumental artifacts such as negative intensity values and extensive sinusoidal structure are to be removed [6, Pg. 427].

5.2 Deconvolution Procedures

Processes for the removal of effects of sidelobes or defects in the dirty images are in fact deconvolution procedures. There exists different methods such as the CLEAN algorithm, MEM standing for Maximum Entropy Method, NNLS for non-negative, least squares algorithm for this report we'll focus on the CLEAN algorithm [for the others see 6, Sec. 10.3].

5.2.1 The CLEAN algorithm

[6, From Pgs. 427-429] The CLEAN algorithm is one of the most successful deconvolution procedures devised by Högbom in 1974. It is basically a numerical deconvolving process applied in the (l, m) domain. It involves breaking down of the intensity distribution into point-source responses and then replacing each one with the corresponding response to a “clean” beam, that is, a beam free of sidelobes, particularly negative ones, and that its Fourier transform should be constant inside the sampled region of the (u, v) plane and rapidly fall to a low level outside it. However these characteristics are essentially incompatible since a sharp cutoff in the (u, v) plane results in oscillations in the (l, m) plane.

Here are the principal steps to follow :

Firstly, we compute the map and the response to a point source by Fourier transformation of the visibility and the weighted transfer function. We refer the functions, synthesised intensity and synthesised beam as the “dirty map” and the “dirty beam,” respectively. We must also ensure that the spacing of the sample points in the (l, m) plane does not exceed about one-third of the synthesised beamwidth. We then find the highest intensity point on the map and subtract the response to a point source, including the full sidelobe pattern, centered on that position assuming that each dirty-beam response subtracted represents the response to a point source. The visibility function of which is a pair of real and imaginary sinusoidal corrugations that extend to infinity in the (u, v)

plane. The peak amplitude of the subtracted point source is equal to γ times the corresponding map amplitude where, γ , is the loop gain, by analogy with negative feedback in electrical systems, and commonly has a value of a few tenths. We record the position and amplitude of the component that is removed by inserting a delta-function component into a model that will become the cleaned map. This process is then repeated iteratively until all significant source structure has been removed from the map. There are several possible indicators of this condition for example, one can compare the highest peak with the rms level of the residual intensity, look for the first time that the rms level fails to decrease when a subtraction is made, or note when significant numbers of negative components start to be removed. These 3 steps can be represented by a model intensity distribution consisting of a series of delta functions with magnitudes and positions representing the subtracted components. Since the modulus of the Fourier transform of each delta function extends uniformly to infinity in the (u, v) plane, the visibility is extrapolated as required beyond the cutoff of the transfer function.

As delta-function components do not constitute a satisfactory model for astronomical purposes. Groups of delta functions with separations no greater than the beamwidth may actually represent extended structure. So as a fourth step, we convolve the delta functions in the cleaned model with a clean-beam response, that is, we replace each delta function with a clean-beam function of corresponding amplitude. The clean beam is often chosen to be a Gaussian with a half-amplitude width equal to that of the original synthesised (dirty) beam, or some similar function that is free from negative values. A Gaussian beam is preferred as it introduces a Gaussian taper in the (u, v) plane which tapers the measured data and the unmeasured data generated by CLEAN and the resulting intensity distribution no longer agrees with the measured visibility data. However, the absence of large, near-in sidelobes improves the dynamic range of the image, that is, it increases the range of intensity over which the structure of the image can reliably be measured. This thereby also removes the danger of over-interpretation. At the point that the component subtraction is stopped, it is generally assumed that the residual intensity distribution consists mainly of the noise. Retaining the residual distribution within the map is, like the convolution with the clean beam, a non ideal procedure that is necessary to prevent misinterpretation of the final result. Finally, we add the residual intensities obtained above into the clean-beam map, which is the output of the process. If we did not add the residuals, there would be an amplitude cut-off in the structure corresponding to the lowest subtracted component. Also, the presence of the background fluctuations provides an indication of the level of uncertainty in the intensity values.

Any intensity feature for which the visibility function is the same within the (u, v) area sampled by the transfer function would produce a response in the map identical to the point source response. CLEAN was initially developed on the basis of the situation pointed out by Högbom that much of the sky is a random distribution of point sources on an empty background. It can also be regarded as an interpolation in the (u, v) plane. Nevertheless, experience shows that CLEAN also works on extended and complicated sources.

The issue we discussed above about this equation 5.2 which was that we cannot directly divide out the weighted transfer function on the right-hand side of equation because it is truncated to zero outside the areas of measurement. In CLEAN, this problem is solved by analysing the measured visibility into sinusoidal visibility components and then removing the truncation so that they extend over the full (u, v) plane. Selecting the highest peak in the (l, m) plane is equivalent to selecting the largest complex sinusoid in the (u, v) plane.

Here following in figures 5.2 is an example of the effect of processing with the CLEAN algorithm.

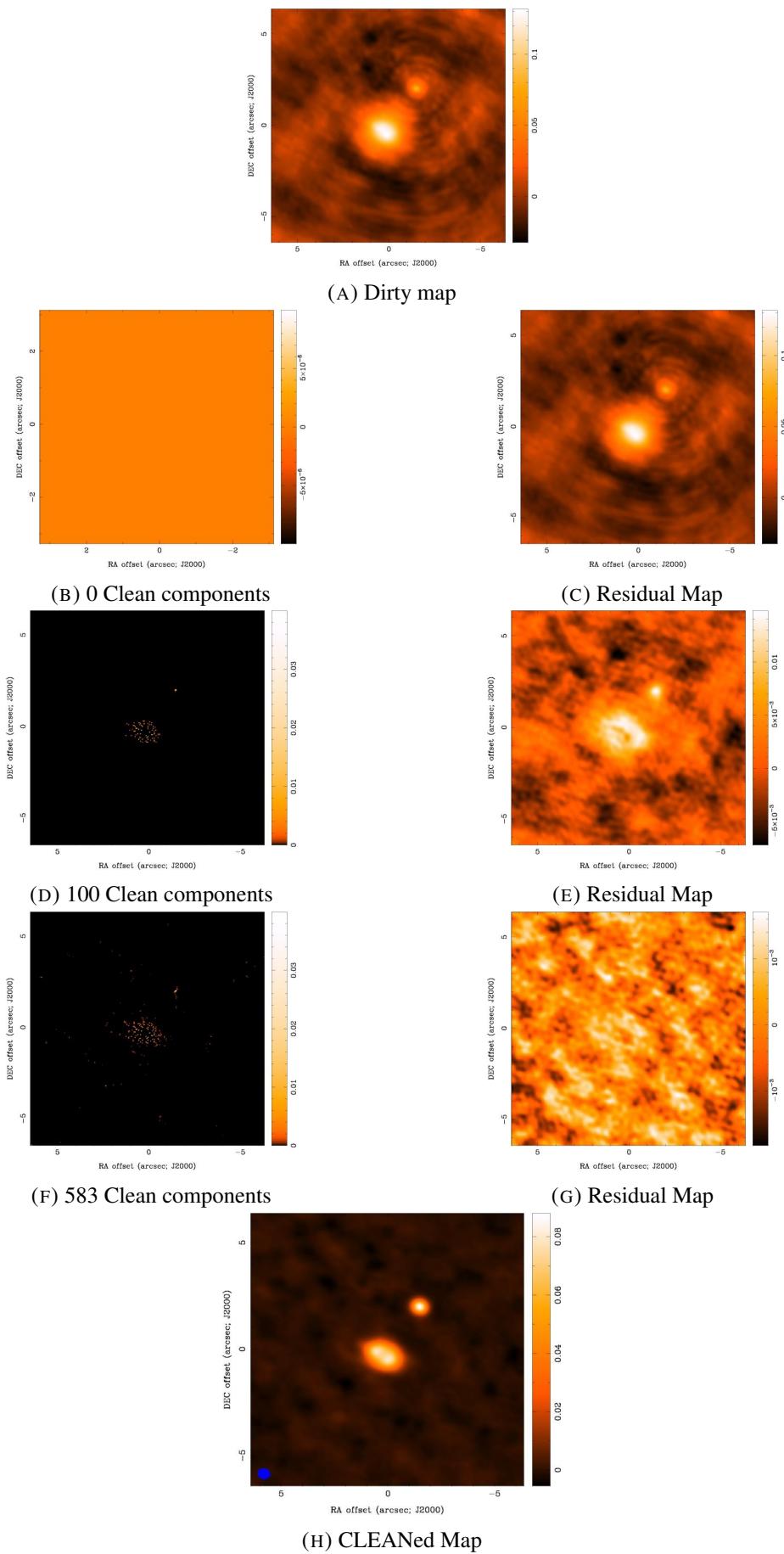


FIGURE 5.2: CLEAN process [1, Slide 54,56,58,59]

5.2.2 Performance of the CLEAN algorithm

[6, From Pgs. 429-432] As a procedure for removing sidelobe responses, CLEAN is easy to understand. Being highly nonlinear, however, CLEAN does not yield readily to a complete mathematical analysis. Some conclusions have been derived by Schwarz in 1978, who has shown that conditions for convergence of CLEAN are that the synthesised beam must be symmetrical and its Fourier transform, that is, the weighted transfer function, must be non-negative. These conditions are fulfilled in the usual synthesis procedure. Schwarz's analysis also indicates that if the number of delta-function components in the CLEAN model does not exceed the number of independent visibility data, CLEAN converges to a solution that is the least-squares fit of the Fourier transforms of the delta-function components to the measured visibility. In enumerating the visibility data, either the real and imaginary parts or the conjugate values (but not both) are counted independently. In maps made using the FFT algorithm there are equal numbers of grid points in the (u, v) and (l, m) planes, but not all (u, v) grid points contain visibility measurements. To maintain the condition for convergence it is a common procedure to apply CLEAN only within a limited area, of the original map.

In order to clean a map of a given dimension, it is necessary to have a beam pattern of twice the map dimensions so that a point source can be subtracted from any location in the map. However, it is often convenient for the map and beam to be the same size. In this case only the central quarter of the map can be properly processed. Thus, it is commonly recommended that the map obtained from the initial Fourier transform should have twice the dimensions required for the final map. As mentioned above, the use of such a window also helps to ensure that the number of components removed does not exceed the number of visibility data and, in the absence of noise, allows the residuals within the window area to approach zero.

Several arbitrary choices influence the result of the CLEAN process. These include the parameter γ , the window area, and the criterion for termination. A value between 0.1 and 0.5 is usually assigned to γ , and it is a matter of general experience that CLEAN responds better to extended structure if the loop gain is in the lower part of this range.

A well-known problem of CLEAN is the generation of spurious structure in the form of spots or ridges as modulation on broad features. The algorithm locates the maximum in the broad feature and removes a point-source component, as illustrated in Fig. 5.3.

The negative sidelobes of the beam add new maxima, which are selected in subsequent cycles, and thus there is a tendency for the component subtraction points to be located at intervals equal to the spacing of the first sidelobe of the synthesised (dirty) beam. The resulting map contains a lumpy artifact introduced by CLEAN, but the map is consistent with the measured visibility data. Cornwell (1983) has introduced a modification of the CLEAN algorithm that is intended to reduce

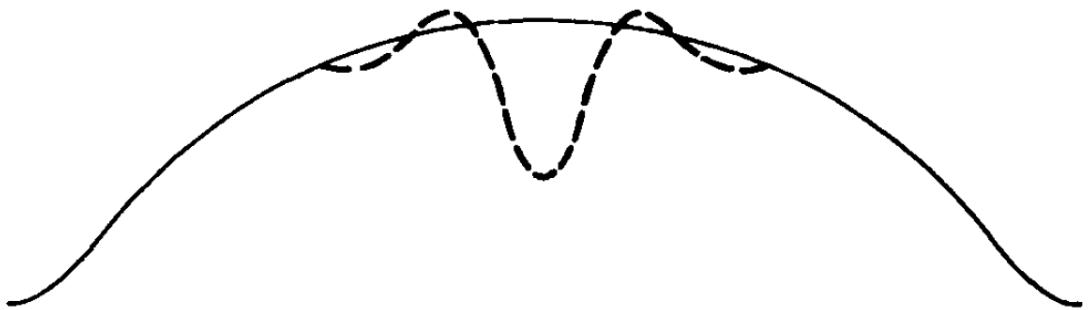


FIGURE 5.3: Point-source response (broken line) removal at the maximum of a broad feature, Clark (1982) [6, Pg. 431, Fig. 11.2]

this unwanted modulation. The original CLEAN algorithm minimizes

$$\sum_i w_i |V_i^{meas} - V_i^{model}|^2 \quad (5.3)$$

where V_i^{meas} the measured visibility at (u_i, v_i) , w_i is the applied weighting, and V_i^{model} the corresponding visibility of the CLEAN-derived model. The summation is taken over the points with nonzero data in the input transformation for the dirty map. Cornwell's algorithm minimizes

$$\sum_i w_i |V_i^{meas} - V_i^{model}|^2 - \kappa s \quad (5.4)$$

where s is a measure of smoothness and κ , is an adjustable parameter. Cornwell finds that the mean-squared intensity of the model, taken with a negative sign, is an effective implementation of s .

The effects of visibility tapering appear in both the original map and the beam, and thus the magnitudes and positions of the components subtracted in the CLEAN process should be largely independent of the taper. However, since tapering reduces the resolution, it is a common practice to use uniform visibility weighting for maps that are processed using CLEAN. Alternatively, in difficult cases such as those involving extended, smooth structure, reduction of sidelobes by tapering may improve the performance of CLEAN.

In 1980, Clark introduced an important reduction in the computation required for CLEAN. This is based on subtraction of the point-source responses in the (u, v) plane and using the FFT for moving data between the (u, v) and (l, m) domains. The procedure consists of minor and major cycles. A series of minor cycles is used to locate the components to be removed by performing approximate subtractions using only a small patch of the synthesised dirty beam that includes the main beam and the major sidelobes. Then in a major cycle the identified point-source responses are subtracted, without approximation, in the (u, v) plane. That is, the convolution of the delta

functions with the dirty beam is performed by multiplying their Fourier transforms. The series of minor and major cycles is then repeated until the required stop condition is reached. Clark devised this technique for use with data from the VLA(Very Large Array (Radiotelescope)) and found that it reduced the computation by a factor of two to ten compared with the original CLEAN algorithm.

To summarize the characteristics of CLEAN, we note that it is simple to understand from a qualitative viewpoint and straightforward to implement, and that its usefulness is well proven. On the other hand, a full analysis of its response is difficult. The response of CLEAN is not unique, and it can produce spurious artifacts. It is sometimes used in conjunction with model-fitting techniques; for example, a disk model can be removed from the image of a planet and the residual intensity processed by CLEAN (Thompson et al. [6, Pg. 432]).

Chapter 6

Conclusions on the report

To conclude this report we might say that the field of Imaging in Radio Astronomy is a non-exhaustive one, it is interesting to appreciate that there are an innumerable set of factors which come into the play and sum up as a whole to produce an end-result. We have approached the subject from the viewpoint of the Signal & Image Processing module where we have focused on physical and extensively mathematical descriptions which would enable one to have a clue on how the data is to be processed, however even in this scope we have certainly not taken the most effective route, there is still much improvement to be done in this report to make something more comprehensive and self-supporting on the subject to avoid unexplained areas, or to reduce the focus to more effective areas in the scope of the module. An improvement could be the inclusion of actual practical examples for the reader to practice with simple generated or raw data by using the acquired techniques and knowledge during the course, so that in this way the unexplained areas mostly concerning the use of particular mathematical functions can be easily grasped. To wrap up we thus encourage the target readers to consult the literature, where everything is extensively explained and the scope of this field does not end at image processing, the broader field of physics behind is something that we encourage the reader to look for.

Appendix A

Appendix Title Here

Write your Appendix content here.

Bibliography

- [1] (Harvard-Smithsonian Center for Astrophysics) David J. Wilner. NRAO: Imaging and deconvolution. https://science.nrao.edu/science/meetings/2014/14th-synthesis-imaging-workshop/lectures-files/wilner_siw2014.pdf, 2014. Accessed: 2014.
- [2] Encyclopedia Britannica. Sir Martin Ryle, 2014. URL <http://www.britannica.com/EBchecked/topic/514574/Sir-Martin-Ryle>. [Online; accessed 2014].
- [3] Encyclopedia Britannica. Quasars, 2014. URL <http://www.britannica.com/EBchecked/topic/486494/quasar>. [Online; accessed 2014].
- [4] James Di Francesco. NAASC: A crash course in radio astronomy and interferometry: Aperture Synthesis. https://science.nrao.edu/science/meetings/ppt-presentation/jdf.webinar.2.pdf/at_download/file, 2014. Accessed: 2014.
- [5] pjbevel@gmail.com. The Fourier Transform of the Box Function, 2014. URL <http://www.thefouriertransform.com/pairs/box.php>. [Online; accessed 2014].
- [6] A Richard Thompson, James M Moran, and George W Swenson Jr. *Interferometry and synthesis in radio astronomy*. John Wiley & Sons, 2008.
- [7] Wikipedia. Dirac comb — wikipedia, the free encyclopedia, 2014. URL http://en.wikipedia.org/w/index.php?title=Dirac_comb&oldid=619618835. [Online; accessed 2014].
- [8] A. Woods, Michael Inggs, and Alan Langman. Accelerating software radio astronomy fx correlation with gpu and fpga co-processors. 2010. URL http://www.rrsg.uct.ac.za/members/drew/andrew_woods_msc.pdf.