UNIVERSITY COLLEGE LONDON

Faculty of Mathematics and Physical Sciences

Department of Physics & Astronomy

A NOVEL INSTRUMENT FOR THE ADVANCED INVESTIGATION OF EXOPLANETS.

Thesis submitted for the Degree of Doctor of Philosophy of the University of London

by

Nicolas Clarke

Supervisors:	Examiners:
Professor Bruce Swinyard	.
Dr Giorgio Savini	.

June 29, 2015

To Amanda and Mum.

I, Nicolas Clarke, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis sets out to analyse and define an instrument capable of advancing the field of exoplanet research. This exciting and nascent field of astrophysics is moving beyond simple population counts and into more detailed characterisation of exoplanet parameters. To this end high resolution and high contrast missions are required to achieve this. In this thesis three designs were investigated, two pupil masked telescopes and an interferometric system.

The First masked design was an axially symmetric system using a petal mask. This design was intended to reduce the diffraction wings resulting from a circular aperture by apodising the pupil. After analysis in software it was found to be insufficient in reducing the airy rings to an acceptable level by a factor of 10^5 and a new design was sought.

Secondly a non axially symmetric mask was tried based on the spergel mask design. This displaced the light from an on axis source away from two detection zones allowing a faint off axis companion to be registered at much higher contrast. Again however the resulting contrast was insufficient to meet science goals.

The third design used was an interferometer based on the Intensity Interferometer model of Hanbury Brown and Twiss. The resolution, $< 0.5mas$, and time taken per observation, $< 3600s$ for a signal to noise of 5, of a moderately sized system was found to be within reach. An attempt was made to verify this on in the laboratory.

A space based mission profile that outlines this system is included. This mission uses a sparse aperture architecture of between 3 and 10 free flying satellites.

Acknowledgements

This thesis was a labour of joy, however it would not have been remotely possible if it were not for the assistance and effort of many collaborators, friends and girlfriends.

Firstly I thank my supervisor Professor Bruce Swinyard for his inspiration and ability to make things look easy. Many times when confronted with a problem he was able to produce a knowledgeable and simple solution. His dedication and assistance in spite of his own challenges was hugely appreciated and this thesis would not have been written without his help.

Help and assistance came from many places, Giorgio Savini who stepped in as my main supervisor when required despite his laready significant workload, Peter Doel and David Brooks who were both ready to provide technical assistance. My work in the laboratory was ably assisted by David on many occasions. My thanks go to Theo Ten Brumelaar and Nic Scott for their accommodation and instruction while at CHARA at the Mount Wilson Observatory. Marc Ferlet's deep technical knowledge at RAL was very useful.

Several work colleagues have provided me with help and inspiration; I want to thank Sylvia, Luke and Roser for their friendship, Nicola for his help in the lab, and Patrick, Deej, Rich, George, Dave, Paul and others for their help in the pub.

Lastly I must thank Amanda for making it all seem worthwhile.

Roads? Where we're going we don't need roads

Doc Brown, *Back to the future 2*

The impression you leave in the spirit of the people
is more important than the result

Arsene Wenger

Contents

List of Figures

This page was intentionally left blank

List of Tables

This page was intentionally left blank

Chapter 1

An introduction

As in all academic fields astrophysicists have consistently sought better instruments with which to measure the universe. The field of astrophysics faces two particular challenges; that the items of interest are typically both very faint and very small, at least in the case that we remain Earth bound observers.

Despite attention being aimed at situations of violence and energy unfathomable on a terrestrial scale, the resultant influence detectable from Earth is very weak. Matched with this problem, and also an explanation of it, are the tremendous distances involved so that talk of physical distances is rendered truly academic. In its place we consider angular size far more often, and still most objects we consider as unresolved point sources.

1.1 Current technology

For those of us interested in the universe at far-infrared wavelengths[1] the problems are exacerbated by the condition that resolving power decreases with inreasing wavelength, summarised in the Rayleigh criterion as

$$\Theta \approx \frac{\lambda}{d}. \tag{1.1}$$

[1] In this thesis the definition of infrared sub bands will be loosely taken as 0.8 to 5μm - near infrared, 5 to 20μm - mid infrared and 20 to 300μm - far infrared.

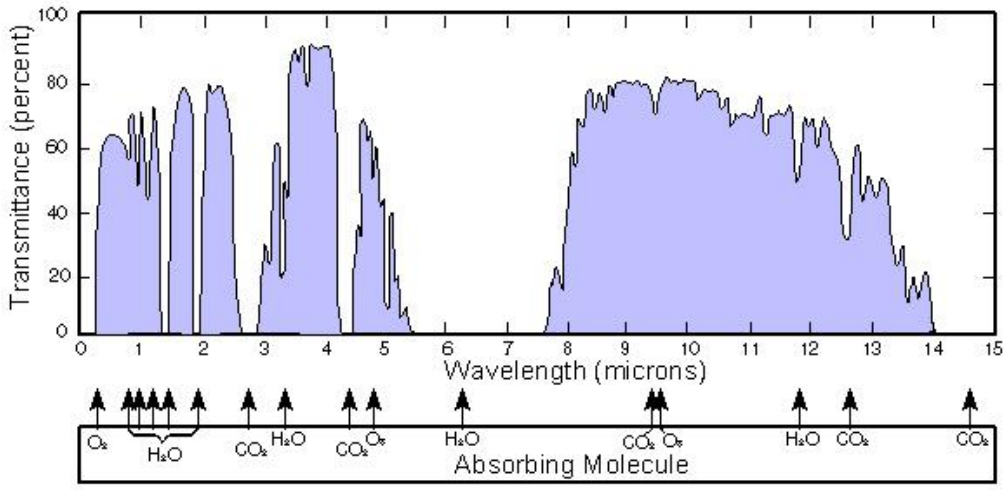

Figure 1.1: Atmospheric transmittance as a function of wavelength showing the windows in near and mid infrared and the cutoff around 40μm. Credit:http:/ewhdbks.mugu.navy.mil/transmit.gif

Hence we either suffer a reduced resolution relative those working at shorter wavelengths or we seek new ways to enlarge our apertures. Additionally Earth's atmosphere has a complex transmittance profile as a function of wavelength with some windows in the near and mid infrared but little to none at wavelengths larger than around 40μm as can be seen in figure 1.1. Several of the molecules that make up the Earth's atmosphere are listed for interest and the absence of these molecules in space are what has driven the production of several infrared satellites. IRAS(Neugebauer et al. 1984), IRTS(Murakami et al. 1996) and ISO(Kessler et al. 1996) paved the way for Spitzer(Werner et al. 2004) and the recently retired Herschel(Pilbratt et al. 2010). The resolution of Herschel with an equivalent 3.28m aperture was around 20 to 100 arcseconds depending on instrument.

In comparison to optical observers for the same size aperture we have the ratio

$$\frac{\Theta_{\mathrm{FIR}}}{\Theta_{\mathrm{Opt}}} = \frac{300\mu\mathrm{m}}{0.5\mu\mathrm{m}}$$

so that $\Theta_{\mathrm{FIR}} \approx 600\Theta_{\mathrm{Optical}}$ and therefore for the same angular resolution we would have to increase our aperture to the kilometre scale relative to the largest optical telescopes in operation. In comparison with radio astronomers we find the ratio closer to

$$\frac{\Theta_{\mathrm{Rad}}}{\Theta_{\mathrm{Opt}}} = \frac{300\mu\mathrm{m}}{10^2\mathrm{m}}$$

and so $\Theta_{\mathrm{Radio}} \approx 10^6\Theta_{\mathrm{FIR}}$. However radio techniques allow for baselines that stretch across

the Earth and therefore they can match the resolving power of a 10 metre FIR instrument with arrays such as the Very Long Baseline Array[2] with baselines in the realm of 10^3km.

So it is these two primary challenges that astronomical instrumentation scientists must combat to push forward the boundaries of the science of astrophysics. Luckily there is a straightforward answer to both lack of flux, the sensitivity problem and tiny angular size, the resolution problem: make it bigger. A larger telescope aperture gathers both more light and more angular information, hence we see a progression of ever larger telescopes from mere centimetres to today's ten metres and the near future of 30 and 40 metre mirrors.

Putting numbers to these challenges gives us a feeling of the challenges ahead, accessible angular resolution with today's facilities are in the micro-arcsecond regime. GRAVITY(Gillessen et al. 2010) to be installed on the Very Large Telescope Interferometer (VLTI)(Glindemann et al. 2003) at the European Southern Observatory will have an astrometric precision of around 10 micro-arcseconds. This is enough for the accurate determination of the black hole Sgr A* at the centre of the Milky Way, giving us clues to the formation and development of our galaxy. Measurements with the MIRC[3] installed at the CHARA(McAlister et al. 2000) interferometer on Mount Wilson have given us images of the gravity darkening of rapid rotators(Monnier et al. 2007), challenging and amending the coefficients of the Von Zeipel laws(de Souza Jr and de Souza Jr 2002). The angular diameter of the stars measured by MIRC are in the realm of 0.1 - 0.4 milli-arcseconds(Monnier et al. 2006).

Single aperture telescopes such as the Keck pair(Colavita et al. 1998) on Mauna Kea and of course Hubble in orbit have imaging capabilities, boosted by adaptive optics in the ground based Keck telescopes, in the region of 0.5 to 1 milli-arcsecond working in the optical and near infrared(Colavita et al. 2003).

Additional to the large cost of such schemes is the impossibility of launching such giant apertures into orbit. Current monolithic schemes are limited by the lifting capabilities of modern launch systems. These restrict the size of any space based telescope to around 3.5 metres as witnessed by the size of Herschel's primary mirror and are unlikely to increase by more than a factor of two within the near future. NASA's Space Launch System only proposes a 5 metre capability. With a segmented mirror the James Webb Space Telescope (JWST)(Clampin 2011)[4] can boast a 6.5 metre primary but again little improvement on

[2]VLBA (Napier et al. 1994)
[3]Michigan Infrared Camera (Monnier et al. 2004)
[4]For a collection of white papers see: http://www.stsci.edu/jwst/doc-archive/white-papers/.

this can be foreseen in the near future.

It is in this light that I have worked on schemes and ideas that take our ability to measure angular resolution to even finer limits. For while we have been building ever larger fully filled apertures the engineering restrictions incumbent upon such schemes means we cannot expect significant development beyond 50 metre apertures in the near future. And telescopes of that size are unlikely to number more than a handful. The costs of the E-ELT[5] are of the order of billions of pounds(Stepp et al. 2003), we cannot rely on such large funding being available for the construction of many of such instruments. Therefore low cost schemes unencumbered by large scale civil engineering requirements are the obvious choice for future development and investment.

1.2 Moving beyond state of the art

Along with the atmospheric problems mentioned above moving a future observatory into space releases the potential for 100 to 1000 metre baselines[6] that offer a flexibility and resolving power that can dwarf anything on the Earth's surface while working in the FIR regime. While we can consider fantastical schemes with baselines stretching to the orbit of Jupiter at 10^6km more conservative approaches have been considered by the likes of Labeyrie(Labeyrie 1996) with an array stretching 10^2km. Darwin(Fridlund 2000) was a nulling interferometer with a baseline of 100m proposal from the European Space Agency (ESA) that lost funding support around 2010 and NASA's SPIRIT concept with a choice between a 30m or 1Km baseline also lost funding at the same time. Recently FIRI, A Far Infrared Interferometer(Helmich and Ivison 2009), also with a 1Km proposed baseline failed to gain support for further study.

This lack of support for both free flying and tethered missions at the current time suggests a serious concern around the feasibility of an interferometer freed from the Earth's atmosphere. It is possible to make many speculations as to why no interferometric mission has flown in the 50 years of the space age but that would be inappropriate in a focussed technical study. One area I will comment on though is the mechanical finesse that the instrument must achieve to carry out its goals. Interferometers promise so much in terms of resolution over single aperture telescopes but are very tricky to build and run. The

[5] European Extremely Large Telescope(Gilmozzi and Spyromilio 2008)

[6] As will be shown in section 2.2 we may in optical terms consider "baseline length" analogous to "aperture diameter" when discussing resolving power of an instrument.

scientific output of orthodox[7] interferometers in the world is still only around 50 papers per year in 2011(Ridgway et al. 2011). As mentioned much speculation for this is possible but again I shall refer to the mechanical difficulties of operating an interferometric observatory.

At CHARA, perhaps the most productive interferometric observatory operating today, they operate a maximum of six apertures with the light coming to a central beam combining laboratory. The optical path from front optical surface to front optical surface of an instrument involves 20 surfaces. Roughly a further 10 surfaces are present in the instrument before reaching a detector. Maintaining a very complex instrument is obviously very taxing and moving such an arrangement into space beyond the reach of maintenance or adjustment has proved to be too large a risk for the scientific community to consider.

One such mechanical restriction we can calculate is the requirement on optical surface alignment via consideration of the coherence length of the incident light. This is a physical length over which a wave packet maintains some singular personality and it is these packets that are correlated at separate apertures in a working interferometer. This length, l is given by dividing the packet's speed, c, by the frequency of oscillation of the physical phenomena we are correlating, ν,

$$\begin{aligned} l_{\text{orthodox}} =& \frac{c}{\nu}\text{m}, \\ =& \frac{3\cdot 10^8}{10^{13}}\text{m}, \\ \approx& 3\cdot 10^{-5}\text{m} \end{aligned} \tag{1.2}$$

$$\begin{aligned} l_{\text{intensity}} =& \frac{c}{\nu}\text{m}, \\ =& \frac{3\cdot 10^8}{10^{7}}\text{m}, \\ \approx& 3\cdot 10^{-1}\text{m}. \end{aligned} \tag{1.3}$$

These frequencies are for light at 1μm from a thermal source, the difference is explained in section 5.1 We can immediately see a benefit for the non orthodox Intensity Interferometry method with a factor 10^4 relaxation in coherence length. This translates to aligning and holding surfaces, all surfaces in total, to an accuracy of around 10's of microns compared to centimetres, clearly a more realistic proposal. The law of "conservation of difficulty"

[7] In this thesis I use the term "orthodox" to refer to the class of interferometer that relies measuring the correlation of the amplitude of a light wave. This is in contrast to an Intensity Interferometer that relies on measuring the correlation of the intensity(The square of the amplitude.). The theory in chapter 2 and appendix A is applicable to both classes and the difference is described in detail in section 5. Note the sub-classifications of "Fizeau" and "Michelson" can apply to both.

present in all instrumentational work declares there must be problems elsewhere with this scheme and indeed there are difficulties.

The fluctuations measured with a Intensity Interferometer are of much lower magnitude than with an orthodox interferometer, which directly leads to a much lower signal to noise ratio. This translates to a longer observing time for the same science output than if we used the orthodox method. Further consideration of these problems are given in chapter 3 with regard to applying Intensity Interferometer to exoplanetary study.

1.3 A summary of chapters

The theory of interferometry is given in chapter 2 showing how the size and shape of a distant object beyond the resolution of (1.1) can be determined by an interferometer using separate distant apertures. The arguments and details in this section are re-elaborated from the books Principle of Optics(Born and Wolf 1999) and Optical Coherence and Quantum Optics(Mandel and Wolf 1995). In the first section the physical propagation of light from a source towards the observer is given using a simple geometric argument. This argument replicates the effects of the underlying physics, the Van-Cittert Zernike theorem and a brief run through of the mathematics of this theorem are given in appendix A.

In chapter 3 the subject of exoplanets is introduced. Exoplanetary research is an exciting nascent field of study for astronomy barely into its second decade. As of summer 2014 around 1800[8] exoplanets have been discovered and many more are suspected. To properly characterise these objects requires accurate host star size measurements and it is here that interferometry can really aid this scientific area as stellar radii measurements are a relatively simple application. This chapter sets out key scientific goals for any future space mission.

In the following chapters description is made of an instrument prototype that is able to achieve the science goals as set out in 3. The form of this instrument varies through three iterations in an attempt to produce a telescope that is able to fully deliver on the science goal requirements. The first of these iterations is a monolithic telescope fitted with a first order spergel mask so that light from a dim target exoplanet, relative to its host star can be separated for spectrographic analysis. The second iteration attempts to do the same image plane manipulation but with a star petal mask. The axial symmetry allowing for

[8]Data taken from www.exoplanet.eu/catalog/ (Schneider et al. 2011)

a much greater "detection zone" than the asymmetric spergel mask. The third iteration is a radical architectural change to an interferometric imaging telescope. This design can better achieve more of the defined science goals. This interferometer design makes use of the "Intensity Interferometer" type operating in a Fizeau configuration. While the feasibility of the first two iterations are calculated in software, a lab top demonstrator was constructed and tested. Additional theoretical explanation is given for this technology ahead of a description of the physical model.

Chapter 4 is a discussion of a project to simulate a masked aperture using both a gaussian petal mask and a Spergel shaped pupil. Included is a description of the code written with Matlab and its usage to draw and evaluate mask designs. Aperture masks are a way to manipulate the electric field entering the instrument in such a way that the image produced is "enhanced" in some way. This can include an increase in resolution beyond that expected from a given aperture size or even to spatially alter the light in the image plane relative to field angle to improve the contrast of desired objects. The ability to characterise the planets via spectroscopy of their atmospheres via spatially resolved imaging is also described.

In chapter 5 a detailed description of an updated lab based Intensity Interferometer based on the work of Hanbury Brown and Twiss is given. An Intensity Interferometer is a special type of interferometer that has seen little application and deployment despite proving its worth in the 1970's at the Narrabri observatory, Australia(Hanbury Brown 1974). Its primary benefit is that it has a much lower sensitivity to mechanical imperfections, and thusly building a large scale instrument is significantly easier than a comparative "orthodox" interferometer. The theoretical background is also given to explain how the advantages come about as worked on by Robert Hanbury Brown and Richard Twiss in the 1950's. This work therefore acts as a proof of concept for the proposals in chapter 3. A space mission proposal of a satellite based Intensity Interferometer using the work of Klein, Guelman and Lipson(Klein et al. 2007) is outlined as a best attempt to reach the constraints set out in chapter 3.

Finally in chapter 6 conclusions are drawn on the general area of this PhD study. Reference is made to scientific goals and how the instrument studied here can aid scientists in making the next generation of discoveries and advancements in astrophysics.

This page was intentionally left blank

Chapter 2

The theory of interferometry

Interferometry is a subsection of the field of physical optics that investigates the correlation of electric fields and is used by astronomers to gather spectral and spatial information of some source distribution. In particular the analysis of the electric field can be described by the "partial coherence function". A helpful aspect of this function is that we can use the concepts of coherence to mathematically describe the physical electric field using readily measurable quantities. The framework uses correlation functions and time averaged intensities which are available to the practical astronomer with "day to day" equipment. This is unlike the actual field variables which oscillate at a frequency that is unreachable with standard laboratory tools; The details of this are given in appendix A, in this chapter a simpler, more applied description is given.

In the venerable status of interferometry there are many full theoretical descriptions of the fundamental physical underpinning the astronomical applications. Complete descriptions can be found in the books Optical Coherence and Quantum optics, Mandel and Wolf, Cambridge(Mandel and Wolf 1995), Cambridge and Principles of Optics, Born and Wolf, Cambridge(Born and Wolf 1999). This chapter will try to summarise the pertinent points required to understand how multiple apertures can be used to investigate spectral and spatial information of some target of interest.

We start with the situation that some source radiates light towards our observatory, it is intercepted in an aperture plane and from there directed at some detector plane for investigation. To simplify we can split this into two physical sections, firstly the space

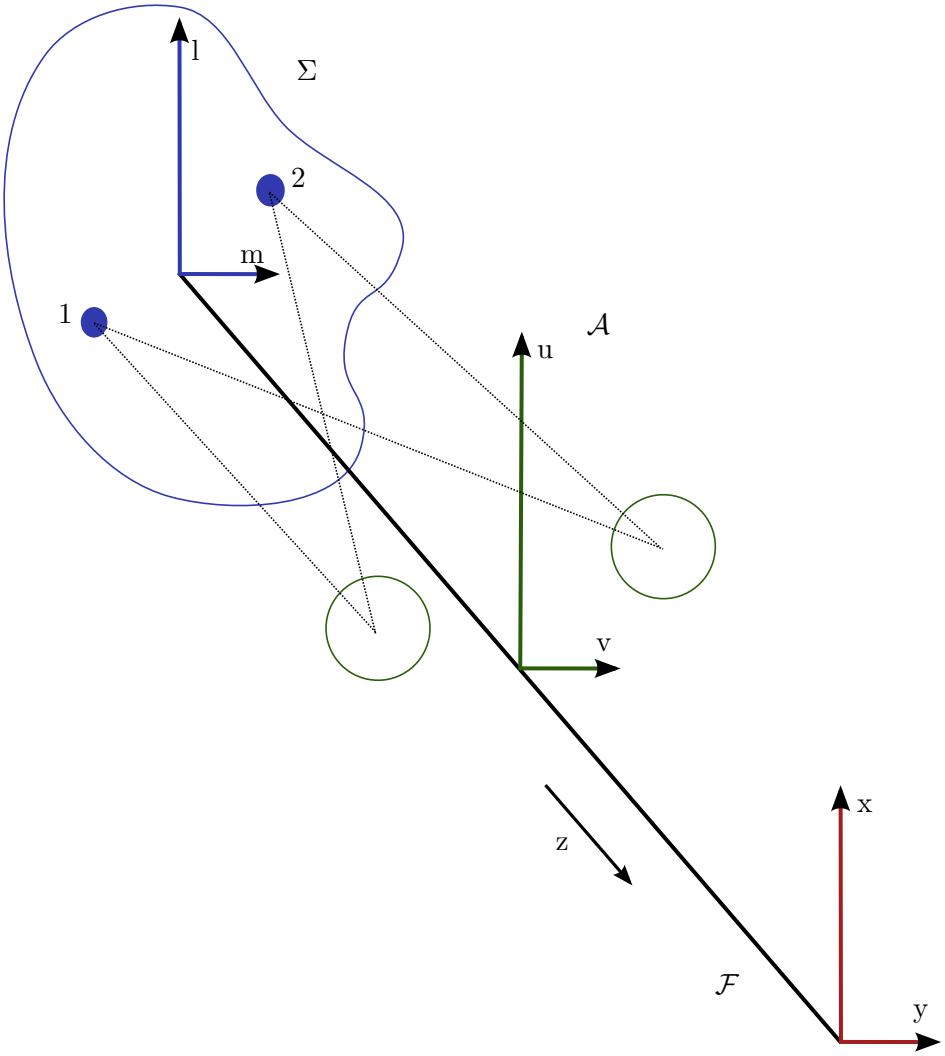

Figure 2.1: Light from a source σ is radiated towards Earth and is intercepted in an aperture plane $\mathcal{A}$. From here the light is brought together at plane $\mathcal{F}$.

from the source to the intermediary aperture plane and secondly from the aperture plane to the detector plane. This situation is shown in figure 2.1

The three planes are thus denoted as the source plane, Σ; the aperture plane, $\mathcal{A}$; and the detector plane $\mathcal{F}$. And from here we can refer to the physical propagation of light across the vacuum of space as the transfer from Σ to $\mathcal{A}$ and the transfer from $\mathcal{A}$ to $\mathcal{F}$ representing the mechanical [1] process of interfering the light fields via our interferometer. The first part for monochromatic situations was put on firm physical and mathematical understanding by Van Cittert and Zernike(van CITTERT 1958) and is described in full in appendix A. It was later advanced to the full polychromatic case by Parrent(Parrent Jr 1960). Here the notion of coherence between two points in $\mathcal{A}$ is explained via a simple geometric description

[1] And later electrical, see chapter 5

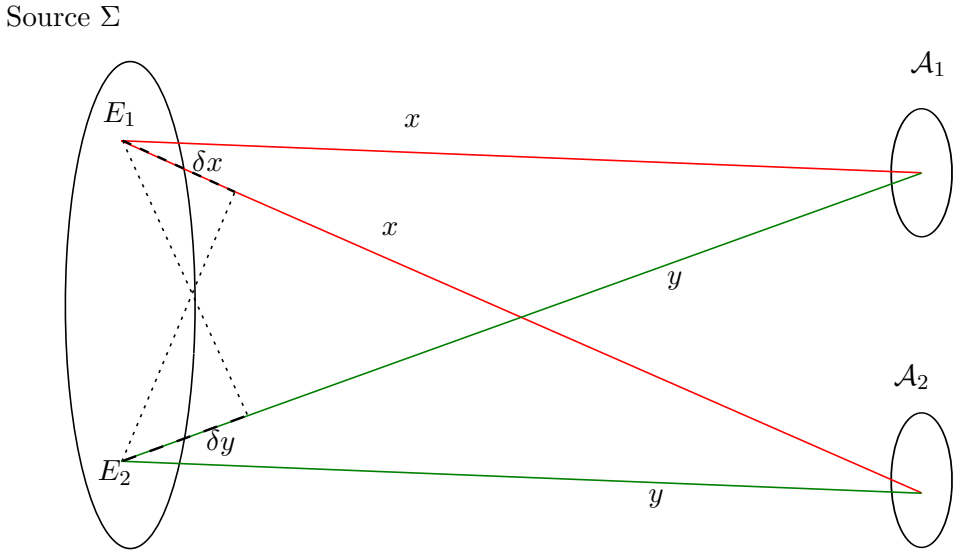

Figure 2.2: A source Σ of angular size θ emits light towards $\mathcal{A}$ from two points E_1 and E_2. The sum of the fields at these two points is of central importance in the theory of interferometry, it is here that coherence is/is not measured.

and this makes up the next section2.1. The second part can be described by reference to the experiment of Young, popularly known as Young's slits and will make up section 2.2.

2.1 A geometric description of coherence

Redrawing figure 2.1 to consider only the propagation from Σ to $\mathcal{A}$ we get figure 2.2. Here we consider the electric field E at E_1 and E_2 and two points in $\mathcal{A}$, $\mathcal{A}_1$ and $\mathcal{A}_2$ that will come to represent the two apertures of a real interferometer. The essence of interferometry is in the fact that while the fields at $\mathcal{A}_1$ and $\mathcal{A}_2$ due to either point alone will be essentially uncorrelated[2] the sum of the fields E_1 and E_2 will be similar at $\mathcal{A}1$ and $\mathcal{A}2$ while the four paths from $E_1$and E_2 to $\mathcal{A}1$ and $\mathcal{A}2$ are of similar length[3].

To state this more mathematically we describe the fields in Σ with a complex number $\tilde{E}(t)$ that fluctuates randomly over time such that the time average is zero, $\left\langle \tilde{E}(t) \right\rangle = 0$. We define the correlation between two fields as Γ

$$\Gamma = \left\langle E_{A1}(t + \frac{\delta x}{c}) E_{A2} * (t) \right\rangle \tag{2.1}$$

and where two fields are perfectly correlated Γ = max and Γ = 0 for perfectly uncorrelated fields.

[2] Assuming the fields to be uncorrelated in Σ

[3] "similar" means less than the coherence length "l" as defined in (2.5).

Given this we can write the summed fields in $\mathcal{A}$ as

$$E_{\mathcal{A}_1}(t) = \langle E_1(x) * E_2(y + \delta y) \rangle \tag{2.2}$$

$$E_{\mathcal{A}_2}(t) = \langle E_2(y) * E_2(x + \delta x) \rangle \,. \tag{2.3}$$

From the geometry of figure 2.2 we can see that Γ will vary as the angle θ varies.

$$\Gamma \to 1 \text{ as long as } x \approx x + \delta x \text{ and } y \approx y + \delta y \tag{2.4}$$

The angle is obviously dependent on: the separation of the planes Σ and $\mathcal{A}$, the separation of $\mathcal{A}_1$ and $\mathcal{A}_2$ and the separation of E_1 and E_2. The astrophysics studied in this thesis limits us to being able to alter only $\mathcal{A}_1$ and $\mathcal{A}_2$, the physical distance to objects of interest is too great, and the physical size of the objects obviously not being amenable to alteration. And it is indeed this that is generally practiced in the field, with two, or more, separate telescopes independently collecting the light from the source.

The approximations in equation (2.4) are constrained by the coherence length, l, of the radiation from Σ where the coherence decreases from 1 to 0 as δx goes from 0 to l. l is defined as

$$l = c\Delta\tau \tag{2.5}$$

with c the speed of light and here τ the coherence time. The coherence time is set by the object emitting the radiation.

Sample distributions of Γ normalised to one in the plane $\mathcal{A}$ are as in figure 2.3; Where the Solid line represents the correlation as given by a binary system of point sources, the dashed line represents a circular distribution and the dashed line a circular distribution with a guassian intensity profile. The game of modelling the source from the visibility curve is still partially a work of art, as can be seen very different source profiles can give very similar curves and typically a real observation will only have 3 or so points with which to model the curve on. As such this is a complicated area of research and not discussed further here beyond that necessary in chapter 3.

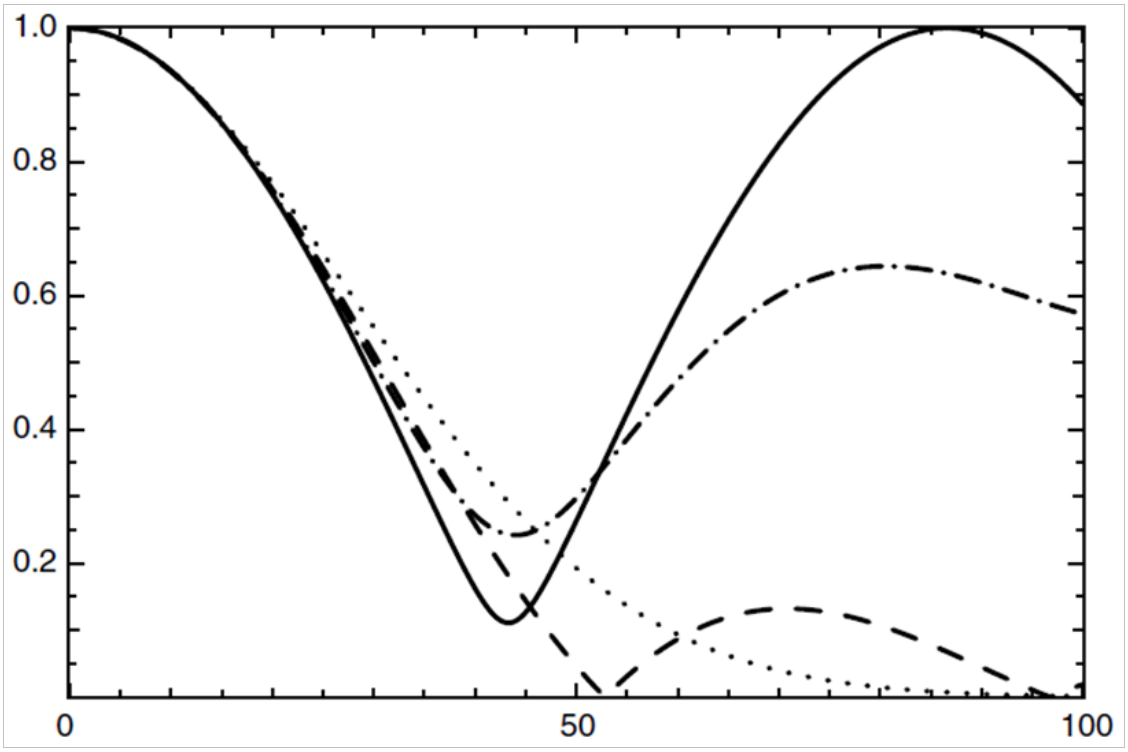

Figure 2.3: Model Visibility Curves. The solid line represents a binary system, the dashed line a flat circular source and the dotted line a circular source with a guassian intensity distribution. Credit:

2.2 Measuring Coherence

As seen in the previous chapter measurement of the coherence in a some plane illuminated by a source can lead to a full description of the intensity distribution in that source. In practical terms it's rarely possible to sufficiently measure the electric field and so some ambiguities reside in any reconstruction of the source. To measure the curves seen in figure 2.3 we can use the experiment of Young where the light from two apertures is brought to focus on a screen and as such is "imaged".

This section represents the second physical setup as mentioned in this chapters introduction where we are considering the passage of the light from $\mathcal{A}$ to $\mathcal{F}$.

After redrawing figure 2.1 again we get figure 2.4 which shows two pinholes $\mathcal{A}_1$ and $\mathcal{A}_2$ in $\mathcal{A}$ allowing light to reach a point f on a screen at $\mathcal{F}$. The light from pinhole 1 to f takes the path r_1 and from pinhole 2 the path r_2. Much like we did in section 2.1 we again look at the electric field disturbance at the point f due to the sum of the disturbances at

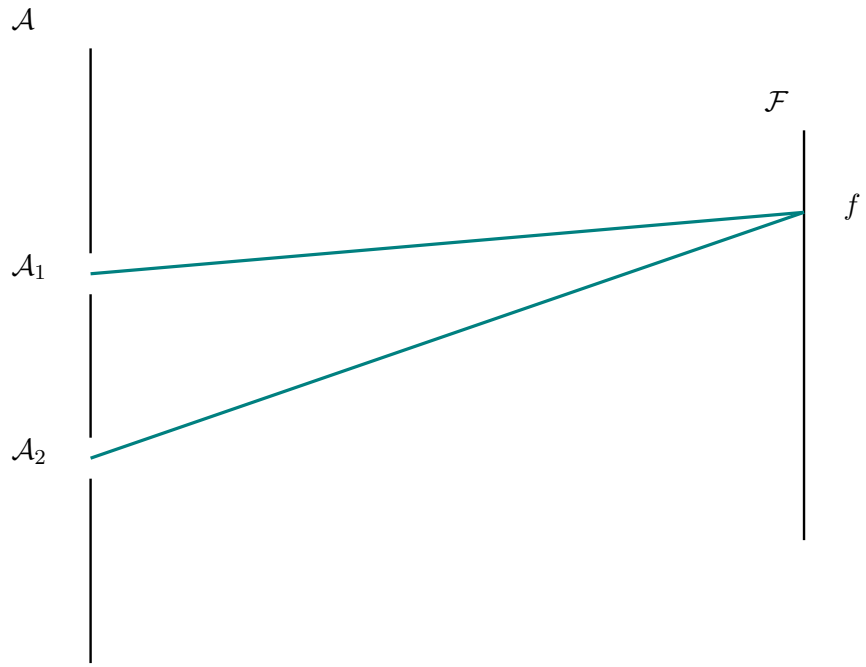

Figure 2.4: Light arriving at the plane $\mathcal{A}$ passes through two aperture $\mathcal{A}_1$ and $\mathcal{A}_2$ and passes to a f in the focal plane $\mathcal{F}$

the two pinholes and write this disturbance as

$$E(f,t) = K_1 E(\mathcal{A}_1, t - t_1) + K_2 E(\mathcal{A}_2, t - t_2), \tag{2.6}$$

t_1 and t_2, are the times taken for the light to get from 1 and 2 to f respectively

$$\begin{aligned} t_1 &= \frac{r_1}{c}, \\ t_2 &= \frac{r_2}{c} \end{aligned} \tag{2.7}$$

and K is factor representing the finite size of the pinholes and is inverse to their size. Taking the intensity I_f measured at f as the time average of the square of $E(f,t)$ we get

$$\begin{aligned} I_f &= E^2(f,t) \\ I_f &= |K_1|^2 I_1 + |K_2|^2 I_2 + 2|K_1 K_2| \Gamma_{12} \left(\frac{r_2 - r_1}{c} \right). \end{aligned} \tag{2.8}$$

The first two terms represent the light reaching f independently from either pinhole 1 or 2. The third term is more interesting in that it takes into account the correlation in the fields at 1 and 2 which as seen previously can vary according to the source distribution. The values of K_1 and K_2 are fixed by the equipment used and so we can let them be absorbed

into I_1 and I_2 respectively without loss of description. Γ is referred to as the mutual coherence function between the fields in $\mathcal{A}_1$ and $\mathcal{A}_2$. When the apertures are co-incident we have the auto-correlation which is physically the same as the intensity,

$$\begin{aligned}\Gamma_{11}(0) &= I_1,\\ \Gamma_{22}(0) &= I_2.\end{aligned} \tag{2.9}$$

At the moment Γ can vary simply due to more light reaching $\mathcal{A}_1$ or $\mathcal{A}_2$ without any change in the actual correlation. The curves in figure (2.4) would look different only for changes in the source brightness and not in shape, and where the source brightness varies because of the intervening transmission this can be misleading. To account for this we normalise Γ to the actual received intensities,

$$\gamma_{12} = \frac{\Gamma_{12}}{\sqrt{I_1}\sqrt{I_2}}. \tag{2.10}$$

where γ is known as the complex degree of coherence.

Finally we recompose the term $\left(\frac{r_2-r_1}{c}\right)$ which represents the phase difference in an exponential form

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)|e^{i[\alpha_{12}(\tau)-2\pi\bar{\nu}\tau]} \tag{2.11}$$

where

$$\alpha_{12}(\tau) = 2\pi\bar{\nu}\tau + \arg\gamma_{12}(\tau). \tag{2.12}$$

With $\bar{\nu}$ used to represent the mean frequency of light.

With these modifications we can write equation (2.8) as

$$I_f = I_{1f} + I_{2f} + 2\sqrt{I_{1f}}\sqrt{I_{2f}}|\gamma_{12}(\tau)|\cos[\alpha_{12}(\tau) - \delta]. \tag{2.13}$$

Here τ and δ are the terms

$$\begin{aligned}\tau &= \frac{r_2 - r_1}{c}\\ \delta = 2\pi\bar{\nu}\tau &= \frac{2\pi}{\bar{\lambda}}(r_2 - r_1),\end{aligned} \tag{2.14}$$

$\bar{\lambda}$ is the mean wavelength.

The behaviour of equation (2.13) on the plane of $\mathcal{F}$ is to produce stripes of light and dark when the cos term is relevant, that is when γ is high, and to produce a low modulation, though of the same frequency, when γ is low. The frequency varies as the difference in the terms τ and δ varies. τ represents the physical size of the path difference whereas δ represents the phase difference over this length.

With this we can reproduce the visibility curves as seen in figure 2.3 with measurements of γ or rather measurements of coherence at different separations. With this interpretation we can now see how, by varying the separation of our apertures we can probe and investigate the coherence of an electric field in $\mathcal{A}$ and from there infer the original source distribution.

Quite how γ is measured depends on the specific instrument and there are various strategies used to maximise the accuracy of the measurement made. In ideal circumstances we might simply scan $\mathcal{F}$ and measure the intensity of the modulation. In more physical circumstances, such as observing through an atmosphere, an additional phase term, ϕ_δ, can be added to δ due to variations in the refractive index along the different paths. This can act to shift the modulation laterally and as such we might be unable to measure the modulation.

To work around this problem an additional path length modulation can be introduced that varies at a frequency higher than that introduced by ϕ_δ. The output from such an operation is to scan across the zero path difference.

Chapter 3

Exoplanets; Their character and characterisation

A fundamental question at the heart of astrophysics is how did our Earth come to support life? And how likely was that? To answer this requires knowledge of the biological processes that come together, these are left for experts in those fields. Prior to the biological activity must have come the geological and prior to this the astrophysical actions of planet formation. And so to understand some of the grandest questions scientists can ask we need to look at how planets form and what populations have been created.

In this very young science we are still working predominantly in the world of modelling and with the limited explanatory power of our solar system. To be able to push the science forwards we need more data, the numbers of discovered exoplanets currently sits at near 2000[1]. And this sample is heavily biased towards the large and the bright members of the family, hot jupiters, massive gas giants sitting very close to their host star. A description of the types is given in section 3.1.

The major method of exoplanetary discovery has either been radial velocity observations or via light curve analysis of transiting planets. The radial velocity (RV) method works on observing the red- and blue-shifting of the host star's spectra as the planet orbits it. As the planet moves towards Earth it will gravitationally tug the star towards

[1]Data for this chapter was taken from www.exoplanet.eu maintained by the Exoplanet Team at the Observatoire de Paris. Where plots are credited to www.exoplanet.eu they were produced with the online plotter tool at http://www.exoplanet.eu/diagrams/.

the Earth as well, lending a blue shift to the spectra. As the planet moves away from the Earth the opposite effect occurs and a red shift is seen. This is shown diagrammatically in figure 3.1. Obviously the exoplanetary orbit is not constrained to sit on a plane with the host star and Earth and therefore a free parameter is introduced that represents the inclination of the orbit to this plane. The timing of the spectral shifts and subsequent velocities measured can both lead to information about the exoplanet. Directly the ratio of the stellar mass, $(m_{\oplus}/m_{\star}) * \sin i$ is given from the velocity and knowledge of $M_{\star}$ gives the planetary mass as a function of i. For a change in inclination of 30°, $\sin(i)$ changes by a factor of 2. Given the geometry of the situation we can expect 87% of exoplanets to lie within 60° and therefore we can be confident that we are at most our planetary mass estimates are within a factor of 2 independently of the actual obital inclination. The size of the orbit is also available from this orbit via the timing of the spectral shifting.

Observations via the transiting method (TM) work on a different premise but one that also relies on identifying the orbital periodicity of an exoplanet. In the case of TM however it is a dip in the apparent brightness of the host star that is observed. This dip is a direct effect of the planet shadowing the stellar light. A typical light curve is given in figure 3.2. again several orbital parameters can be deduced from this data including the radius of the exoplanet, its mass and the orbital radius. Via spectral subtraction an estimation of the composition of the planetary atmosphere can be made. Fuller descriptions of this are given in the following sections.

3.1 Discovering exoplanets

3.1.1 The Radial Velocity method

The RV method makes use of the doppler shifting of the light from the host star, this doppler shift can be converted into a line of sight velocity $V_{\star}$. A plot of this velocity over time allows a sinusoidal curve to be plotted as in figure 3.1. Making use of Kepler's third law,

$$r^3 = \frac{GM}{4\pi^2} P^2, \tag{3.1}$$

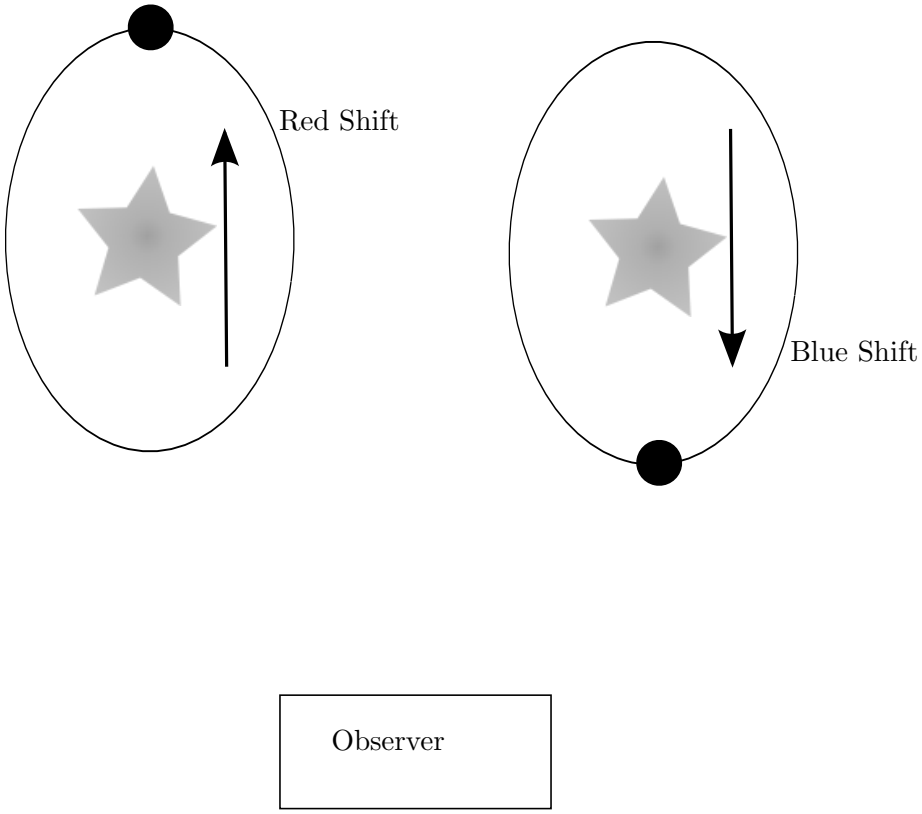

Figure 3.1: On the left the exoplanet's motion is drawing the host star away from the observer and so a red shift is observed in the star's spectra. On the right, after half an orbital period the host star's motion is toward the observer and a blue shift is observed.

where r is the planetary radius, G is Newton's gravitational constant, P is the orbital period and $M \approx m_\star$, the host star mass. The velocity of the planet can be found from

$$v_\oplus = \sqrt{G m_\star v_\star}. \tag{3.2}$$

Here $v_\oplus$ and $v_\star$ are planetary and host star velocities, and while the velocity of the exoplanet is not directly interesting in does lead via a conservation of momentum to the planetary mass $m_\oplus$

$$\begin{aligned} m_\oplus v_\oplus &= m_\star v_\star \\ m_\oplus &= \frac{m_\star v_\star}{v_\oplus}. \end{aligned} \tag{3.3}$$

Fischer et al.(Fischer et al. 2014) report that the "Habitable Zone Planet Finder"(Mahadevan et al. 2012) will operate with an ability to measure down to $1 - 3\text{ms}^{-1}$. This instrument operates in the NIR via a fiber fed spectrograph.

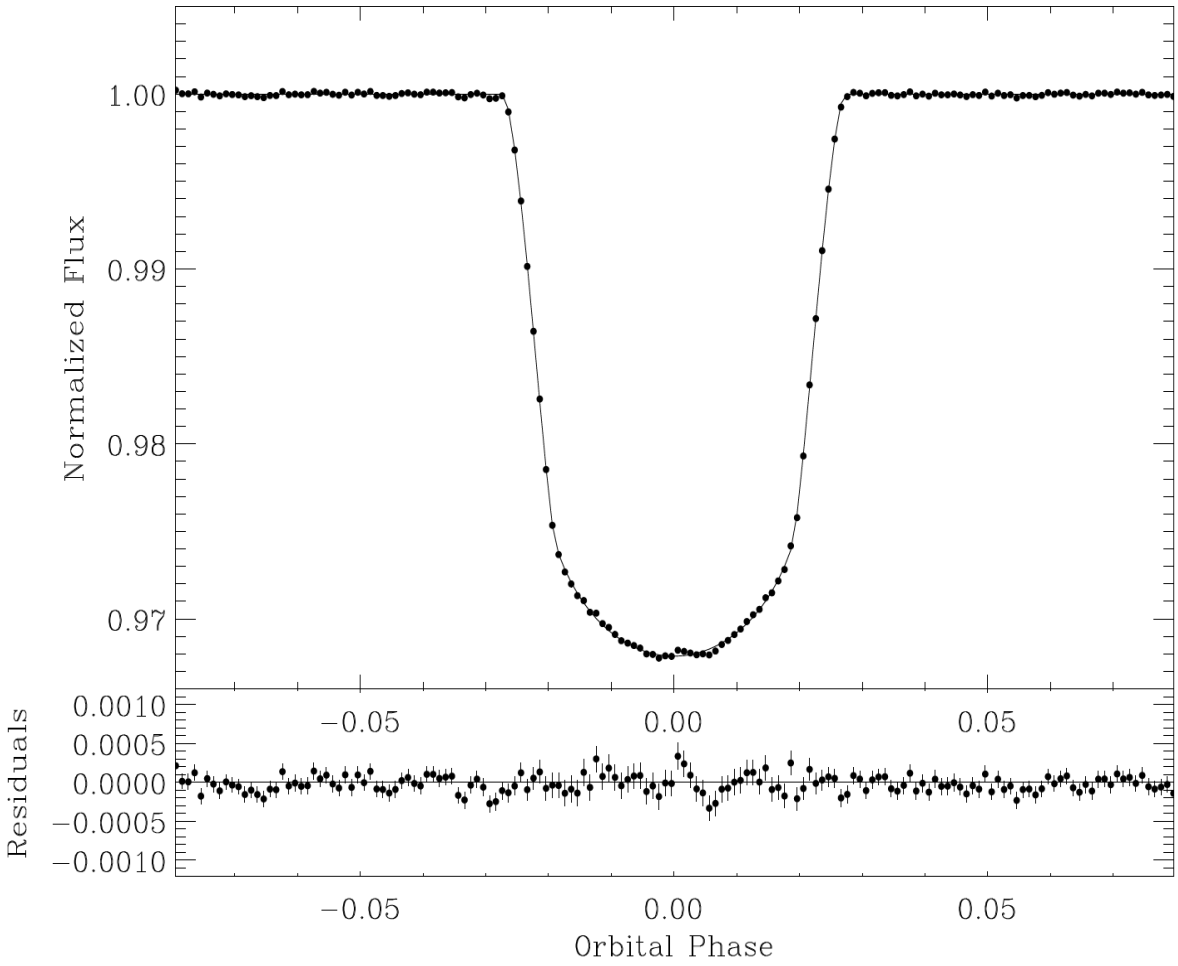

Figure 3.2: As the exoplanet moves in front of its host star it obscures a portion of the stellar light. Typical changes in flux are around 1% of the total light.

3.1.2 The Photometric Transit method

The Photometric Transit method (PT) works on observation of the light curve from a star, this curve, figure 3.2(Credit:(Alonso et al. 2008)), is a plot of flux versus time. In the absence of natural periodic variations the presence of a transiting exoplanet can be deduced from the regular reduction in flux from a star as it is partly obscured by the planetary disc. This obscuration typically leads to a reduction of around 1% of the light from the host star. Usually the curve is measured over several periods and after an estimation of the orbital period P the measurement points are folded into one period. The time when the flux first starts to drop is known as the ingress of the exoplanet and the time as it leaves is the egress, the time between the beginning of ingress and finishing of egress is the transit duration, T. The shape of the curve at these points can give information of the orbital and planetary parameters. As with the RV method detailed above in 3.1.1 we can

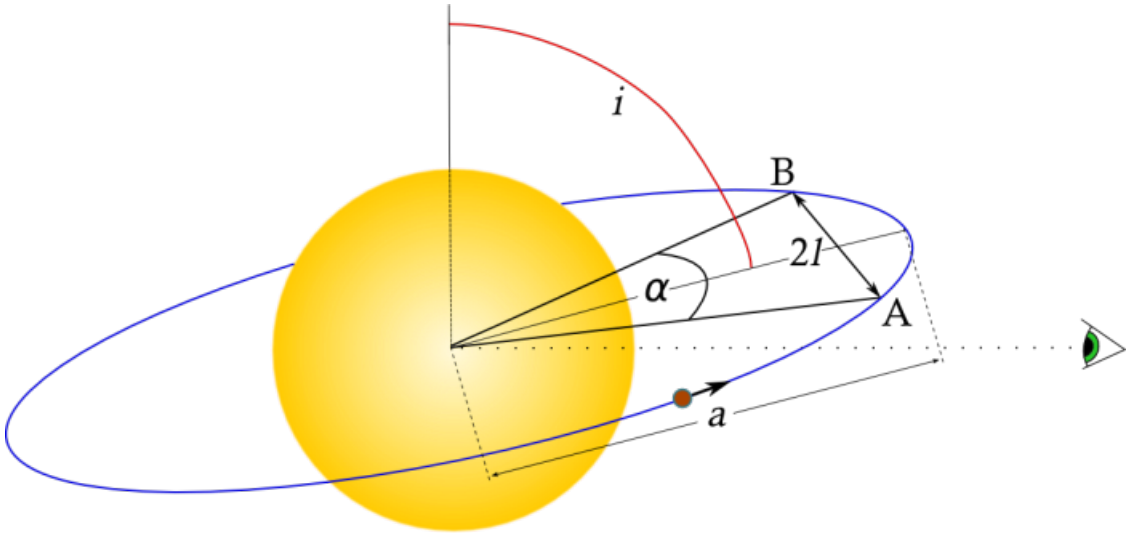

Figure 3.3: The orbit of an exoplanet around its host star can betray several orbital parameters simply through analysis of the resulting light curve. Full description and credit in text.

compute several orbital parameters of the exoplanet via analysis of this light curve. With these we can determine the planetary radius, its orbital inclination and its orbital period.

This method obviously only works for the correct alignment such that the exoplanet obscures part of the stellar disc as seen from Earth. The chances of this happening are roughly equal to the ratio of the planetary radius to the stellar radius. Given typical planetary sizes this leads to a chance of detection of around 3%, thus this method is inherently limited in its ability to find exoplanets. It is however easier to scan a large number of potential stars simultaneously in comparison with the RV method as photometry of a field is easier to do than spectroscopy of a field. Photometry is less photon hungry.

To identify exoplanetary characteristics of radius, orbital radius and mass we need to translate the light curve as seen in figure 3.2 to a physical situation such as seen in figure 3.3[2], Here we see that as the exoplanet traverses from A to B it moves a distance $2l$ on the sky as seen from Earth. a represents the semi major axis of the planets orbit, the inclination of the plane of the orbit i is shown and the portion of the orbit that defines $2l$ is given by α. This geometry leads us to see that the duration T is related to the sum of the radii by

$$T = P\frac{\alpha}{2\pi}, \tag{3.4}$$

[2]credit for both figure 3.3 and figure 3.4:https://www.paulanthonywilson.com/exoplanets/exoplanet-detection-techniques/the-exoplanet-transit-method/.

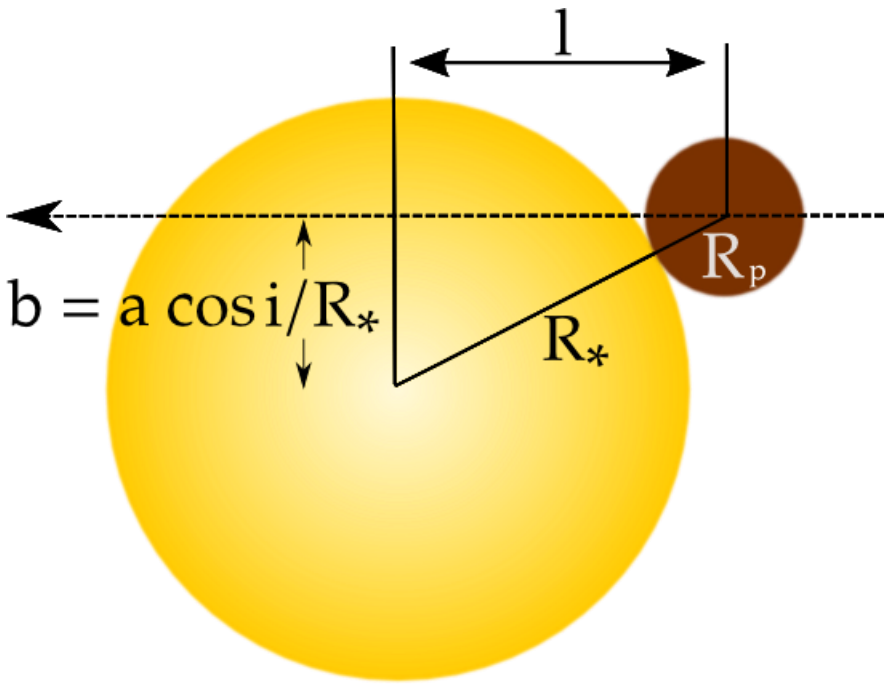

Figure 3.4: Seen in projection the planetary disc travels across the stellar disc. Assuming a circular orbit

where the fraction of the orbit $\alpha/2\pi$ is given by rearranging the triangle A,B and the star's centre to

$$\sin(\alpha/2) = \frac{l}{a}, \tag{3.5}$$

so combining (3.4) and (3.5) we get

$$T = \frac{P}{\pi} \sin^{-1}(l/a). \tag{3.6}$$

l can give us the sum of the planetary and stellar radii. Figure 3.4 and a small amount of pythagoras shows us that $2l$ is equivalent to

$$2l = 2\sqrt{(R_{\text{star}} + R_{\text{planet}})^2 - b^2}. \tag{3.7}$$

So with T and b we can from (3.7) measure the total of the radii and with independent measures of the host star's radius we can calculate the radius of the exoplanet.

Type	Size (Jupiter masses)	Constitution
Mercurian	$10^{-7} - 0.1^{-2}$	Rocky
Sub-Terran	$(0.1 - 0.5) * 10^{-2}$	Rocky
Terran	$(0.5 - 2) * 10^{-2}$	Rocky
Super-Terran	$0.02 - 0.1$	Rocky
Neptunian	$0.1 - 0.5$	Gas Giant
Jovian	> 0.5	Gas Giant

Table 3.1: Table of exoplanet types detailing sizes and constitution.

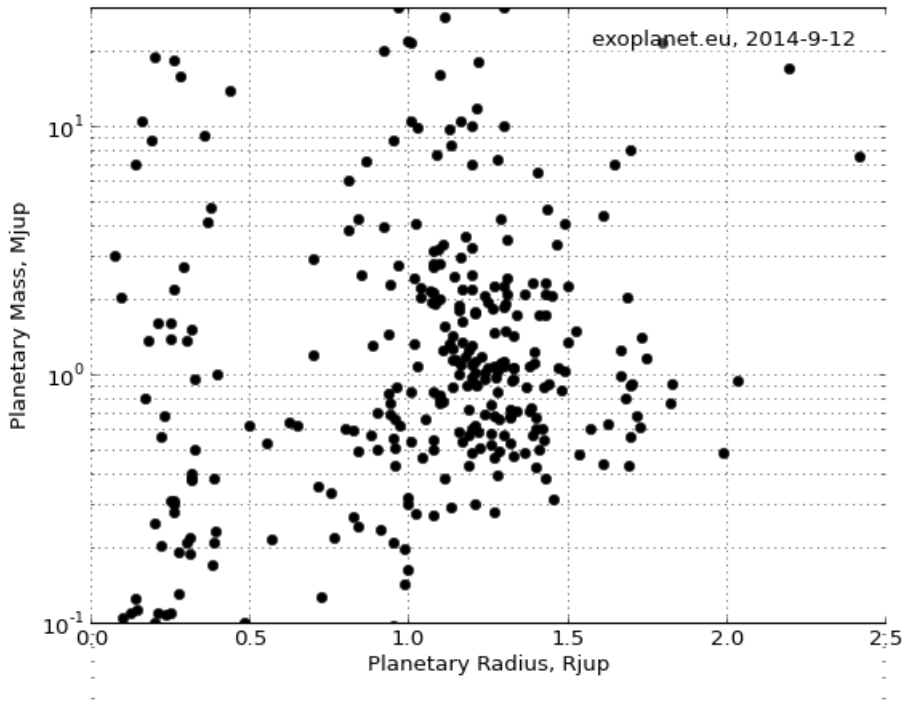

Figure 3.5: Plot of exoplanet mass versus radius . Showing all discovered as of September 2014.

3.2 Current populations of known exoplanets

Exoplanets come in four archetypes, named after analogues to planets in our own solar system. These are, in increasing size, Mercurians, Terrans, Neptunians and Jovians. The first three represent rocky or terrestrial constitutions and the last two gas giants. There sizes are summarised in table 3.1. The current total of discovered and confirmed[3] exoplanets as of September 2014 is 1822 planets in 1137 systems. Figure 3.5 shows the range of

[3]The main source for exoplanet possible detections is NASA's Kepler space mission. This mission has suggested around 4000 stellar candidates as possible host stars, 1822 have been confirmed by further independent ground based observation. This backup confirmation work has had a low false positive rate and so we can expect the number of discoveries to raise significantly over the coming years as more candidates are reassessed.

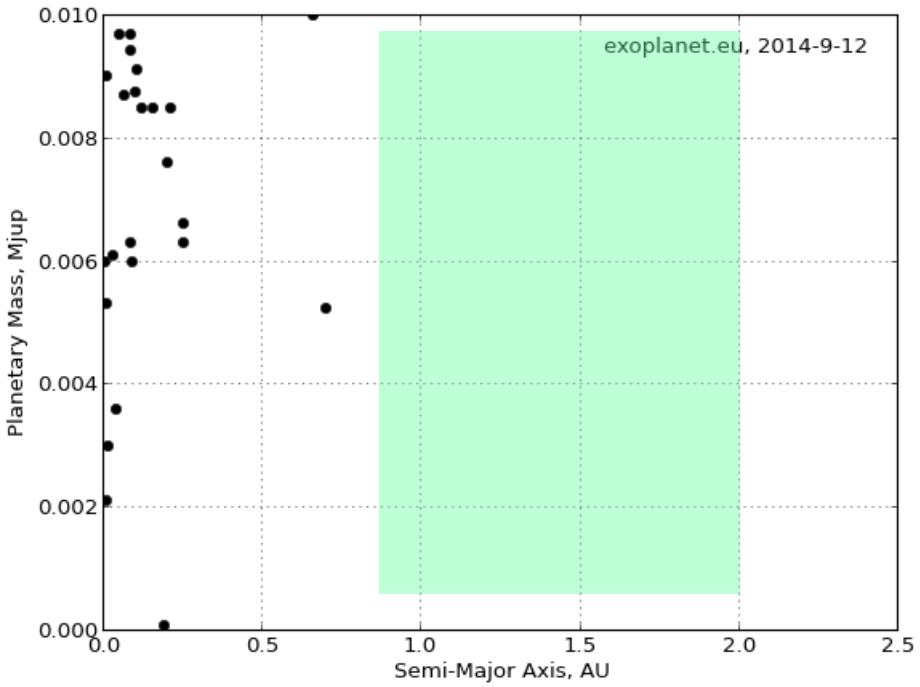

Figure 3.6: Planetary mass in the range $0 - 0.01 m_{\text{Jup}}$ versus orbital semi-major axis in AU. 1 Earth mass $\approx 0.005 m_{\text{Jup}}$. The green area represents the "habitable zone" where we might best expect to find life supporting conditions. The zone is defined in the text.

exoplanetary masses versus radii. These range from around 10^{-1} to just over 10^{1} Jupiter masses (m_{Jup}), and from as small as 10^{-5} Jupiter radii (r_{Jup}) up to 2.4 r_{Jup}. The majority of discoveries lie in a range from 0.3 to 3 m_{Jup} and 0.75 to 1.75 r_{Jup}. As can be seen in table 3.1 this places the discoveries in the Neptunian and Jovian class. Focussing on Terran sized exoplanets gives figure 3.6. Which shows us that almost all the discovered Terran sized planets closest to Earth in size and constitution all sit at an orbital radius of less than 0.3 AU. At this distance the surface temperature is likely to be too high for liquid water and as such not a good candidate for supporting life. This region is known as the "goldilocks zone" or "habitable zone" (HZ). It's definition varies according to author but a sample of recent work suggests[4] it is a region 0.8 to 2 AU from the host star. This is obviously dependent on the luminosity of the host star and the above figures are for a Sol, G type star on the main sequence. This work will concentrate on stars of F,G and K spectral type, cooler less massive stars.

As can be seen in figure3.6 there are currently no discovered planets sitting in the HZ

[4](Traub 2012), (Kaltenegger and Sasselov 2011), (Von Braun et al. 2011)

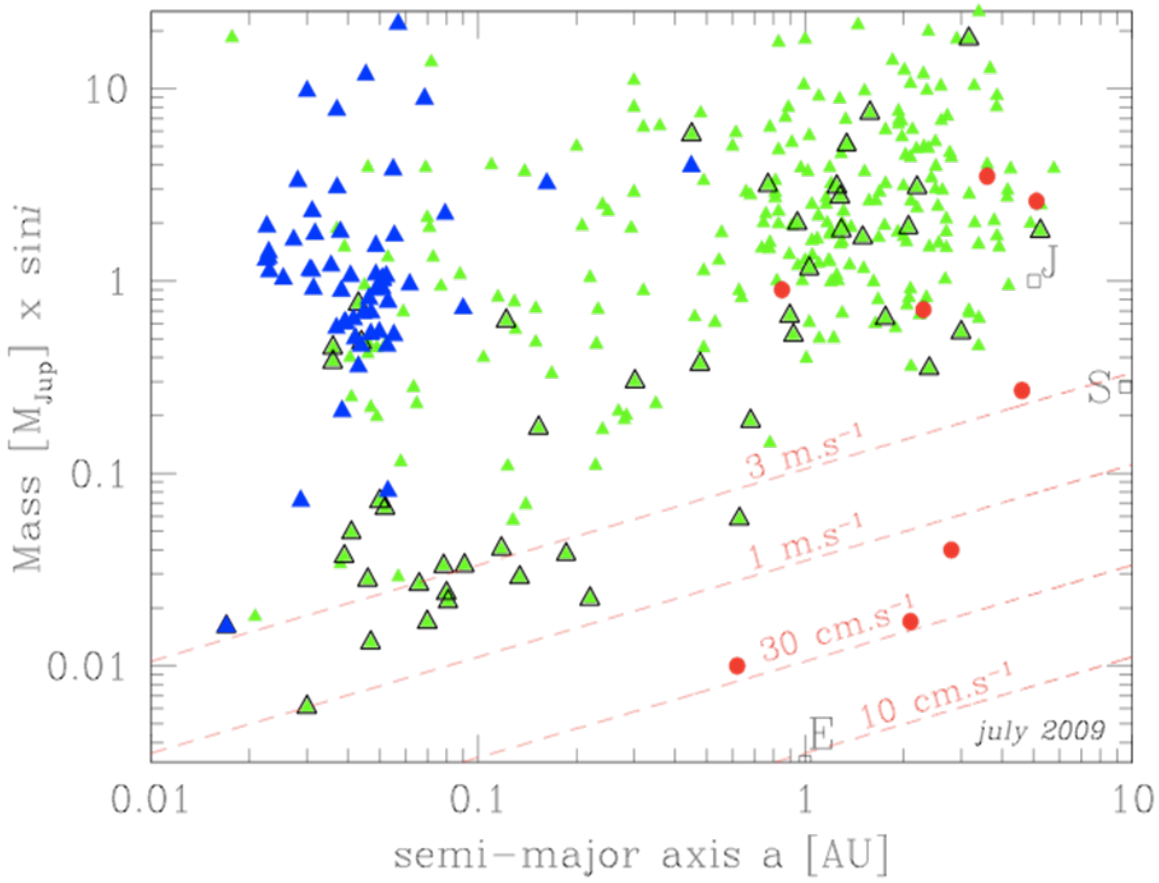

Figure 3.7: Exoplanetary masses are plotted against orbital radius. The dashed lines represents the threshold of potential discovery for a given sensitivity to the RV modulation. This physically represents a velocity of the host star.

with the expected bias towards massive planets that sit closer to there host star. Sitting closer leads to a greater RV effect as does being more massive. Also being more massive has a greater effect on the light curve with respect to the PT method. Figure 3.7(Bouchy et al. 2008) highlights the RV sensitivity against discovered planets in 2009. Sensitivities of less than 1ms^{-}1 are required to find Earth like masses at 1AU. Reading off the plot and taking an Earth like mass to be around $0.5 * 10^{-3} m_{\text{Jup}}$ an RV sensitivity of about 15cms^{-1} would be required at 1AU. This is still a fair way off with regards to current technology, these capabilities are discussed more in 5.4.

Figure 3.8 shows a histogram of discovered exoplanetary masses. There is a bi modal distribution with peaks at $2 m_{\text{Jup}}$ and $2 * 10^{-2}$ representing the Jovian and Terran/Super Terran archetypes. It has been discussed that current survey methods favour discovery of Jovian types but figure 3.8 shows a significant proportion of lower mass planets. This directly implies a higher source population of Terran sized planets to overcome the detection

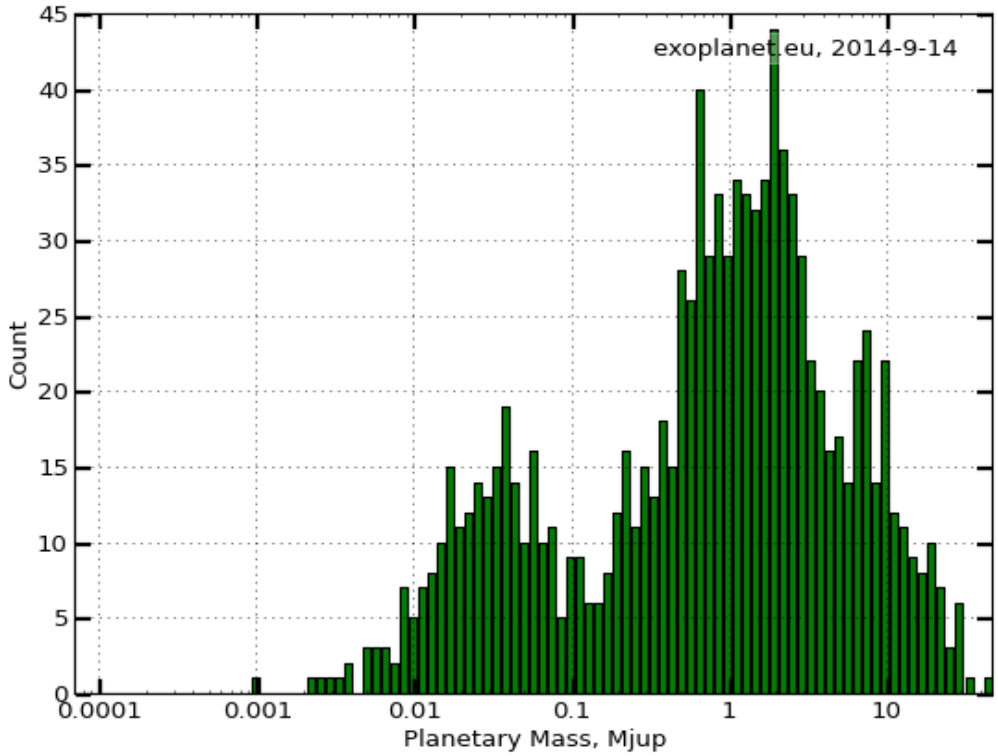

Figure 3.8: Population of exoplanetary masses. Masses are plotted in Jupiter masses on a log scale. Two peaks can be seen, around $1 - 2m_{\mathrm{Jup}}$ and $2 * 10^{-2}m_{\mathrm{Jup}}$.

biases. Foreman-Mackey(Foreman-Mackey et al. 2014) propose an underlying population of 0.02 Earth analogues per star. This is lower than other estimates but with strong (pessimistic) assumptions. Given this abundance they calculate there should be $10.6^{+5.9}_{-4.5}$ transiting Earth like exoplanets in the full, including unconfirmed, Kepler dataset(Borucki et al. 2010).

This is an intriguing result but still only represents a beginning of the study of Terran exoplanets and there is no guarantee that these exoplanets will lie in any HZ around their host stars.

Shifting our focus to look at the HS, we can plot figure 3.9 which shows us planetary mass versus the radius of the host star. There is a clear trend towards stars of close to 1 solar radius (r_{Sol}) with a mean of $1.8r_{\mathrm{Sol}}$ across those stars that have a measurement. Again the source populations of stars in our galaxy[5] will suggest that smaller stars are more common. There is a minimum size of a star at around 10% of the solar radius and mass. So while we might expect a predominance of M class Red Dwarfs, around 76% of

[5]Exoplanetary science is still almost entirely a galactic, not intergalactic enterprise.

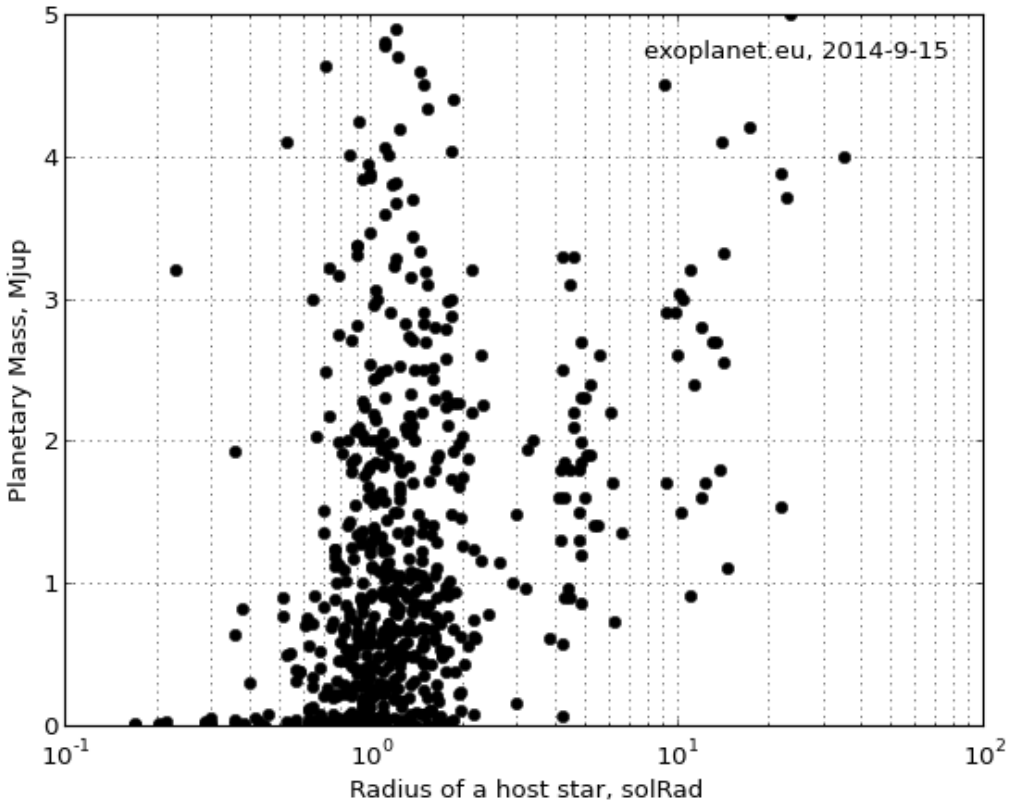

Figure 3.9: Exoplanetary mass plotted against the measured radii of the host star. The plot concentrates on masses $< 5m_{\mathrm{Jup}}$ and we see that stars of $0.7 - 1.1$ solar radius dominate.

stars near our sun are M class stars(Ledrew 2001), in our selection their luminosity drops away to less than 10%(Kaltenegger and Traub 2009) of a solar luminosity. Therefore it is the brighter yet still numerous class that will dominate a survey. The classification of the HS is presented in table 3.2. Concentrating then on these more populous stars we can start to consider the angular sizes we are confronted with when considering these objects. In table 3.3 the angular size in micro-arcseconds of the stellar disc is calculated for typical F,G, K and M class stars at a distance of 40, 60 and 100 parsecs. Even the largest of these classes at a relatively short distance of 40pc is a challenging target of only 302 micro-arcseconds. The more numerous G and K classes at modest distances are of the order of tens of micro-arcseconds. The angular size of 1 AU is in the order of tens of milli-arseconds, a resolution already achieved by both ground based and space based observatories. Considering the habitable zone as seen in figure3.6 a search space of 8 to 100 milli-arcseconds is defined dependent on host star distance. The Gliese star

Stellar class	Count	Percentage
O	0	0
B	10	1
A	12	1
F	116	14
G	370	45
K	243	29
M	77	9

Table 3.2: The surveys are dominated by F,G and K class stars representing 88% of exoplanet hosts. G class stars alone count for 45% of discoveries. Note this table only includes exoplanets for the which the spectral class of the host star is recorded.

Table 3.3: Angular sizes in micro-arcseconds are given for typically sized F, G, K and M class stars at distances of 40, 60 and 100 parsecs. The angular size represented by 1 AU is also included for reference.

Stellar Class	Typical Radius (m) [r_{Sol}]	Size (μ-arcsec) at 40pc	Size at 60pc	Size at 100pc
F	$9.05 * 10^8$ [1.4]	302	101	60
G	$6.96 * 10^8$ [1]	233	78	47
K	$4.18 * 10^8$ [0.6]	140	47	28
M	$1.74 * 10^8$ [0.25]	58	19	12
Also included is 1 AU for comparison, the angular sizes are given in milli-arcseconds				
1 AU	$1.50 * 10^{11}$	50	17	10

catalogue(Gliese 1957) suggests a stellar density of 0.120 stars per cubic parsec out to 5 parsecs. Extrapolating that density to 100 parsecs we find,

$$\text{Starswithin100pc} = 0.120 * \frac{4\pi}{3}(100\text{pc})^3, \tag{3.8}$$

$$= 47557\text{stars}. \tag{3.9}$$

The angular size is not out only concern when considering the difficulties of observing and measuring exoplanets. The flux from the host star relative to the exoplanet is very large. The flux from the planet is almost entirely the reflected flux from the star. Some of the most massive stars are self luminous but these are outside our area of interest which is centred around Terran type exoplanets.

Brown and Burrows(Brown and Burrows 1990) calculate the flux ratio of a star to exoplanet for a Jovian type at an angle most favourable to detection. Their result is

$$\mathcal{R} = \frac{3.1 \times 10^{-8}}{d(\text{pc})^2 \theta(\prime\prime)^2}. \tag{3.10}$$

Where $\mathcal{R}$ is the ratio of flux received at Earth of that produced by the HS and that

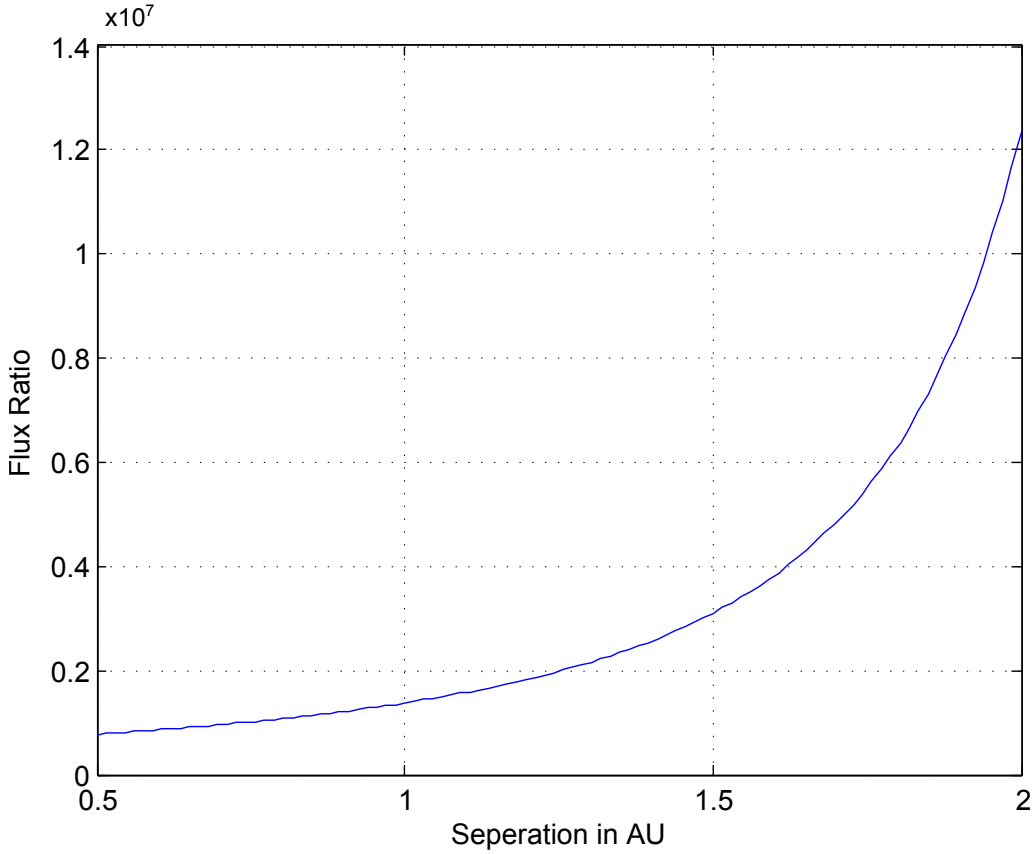

Figure 3.10: Flux ratio of a Jovian sattelite around a G class star. The ratio is plotted as a function of the seperation of the exoplanet and host star.

reflected by the exoplanet, d is the distance to the system from Earth and θ is the angular separation on the sky of the exoplanet and host star in arcseconds. Figure 3.10 plots equation (3.10) to show the variation in flux ratio between the host star and an exoplanet within the bounds of the habitable zone. It can be seen to vary between $0.2 - 1.2 \times 10^{-7}$. This is a challenging setup for equipment and methods to address this problem are given in chapter 4.

This leaves us with some numerical constraints and goals for any future instrument that might wish to able to detect Earth like exoplanets around nearby stars. These can be summarised as follows:

With the data in table 3.4 we can constrain and inform a space mission aimed at the goal of furthering investigation of possible life supporting exoplanets and this is discussed in the following section.

Parameter	Upper Bound	Lower Bound
Reflected flux ratio $\mathcal{R}$	$a = 0.5$AU	$a = 2$AU
$\lambda -$ independent	0.2×10^{-7}	1.2×10^{-7}
Angular size θ	$d = 20$pc	$d = 100$pc
semi-major axis, a	50	10
Stellar radius, R	0.47	0.094
Planetary radius, r	0.005	0.001

Table 3.4: Summary of the limits imposed on Earth bound observers of nearby exoplanets seeking Earth like Terran type planets in the habitable zone. Angular sizes are given in milli-arcsecs.

Chapter 4

Image plane manipulation via pupil masks

Two initial versions of an instrument to reach the goals set out in chapter 3. First a short description of the theory required to understand aperture masking is made before discussion of an apodised pupil mask with plots of resulting diffraction ring reduction. Following this a different attempt to enhance image plane contrast is made using a non axially symmetric Spergel mask. the image plane is plotted for an off axis point source of much lower brightness than an on axis source.

The analytic method used is to draw an x, y map in some array structure where the value of each element represents the electric field at that "point". In reality the code is non-analytic and uses discrete elements to describe both the physical structure of the entrance pupil and the electric field is chopped into finite pieces so that the field can be digitised. Full use is made of Fourier optics such that the "image" output is considered to be in the far field Fraunhofer regime as defined in equation 4.1. This assumption allows for the simplification that the image is therefore simply the Fourier transform of the entrance pupil. The entrance pupil sits in a plane that is equally considered in the far field from the object such that again the electric field in the aperture plane is the Fourier transform of the object plane.

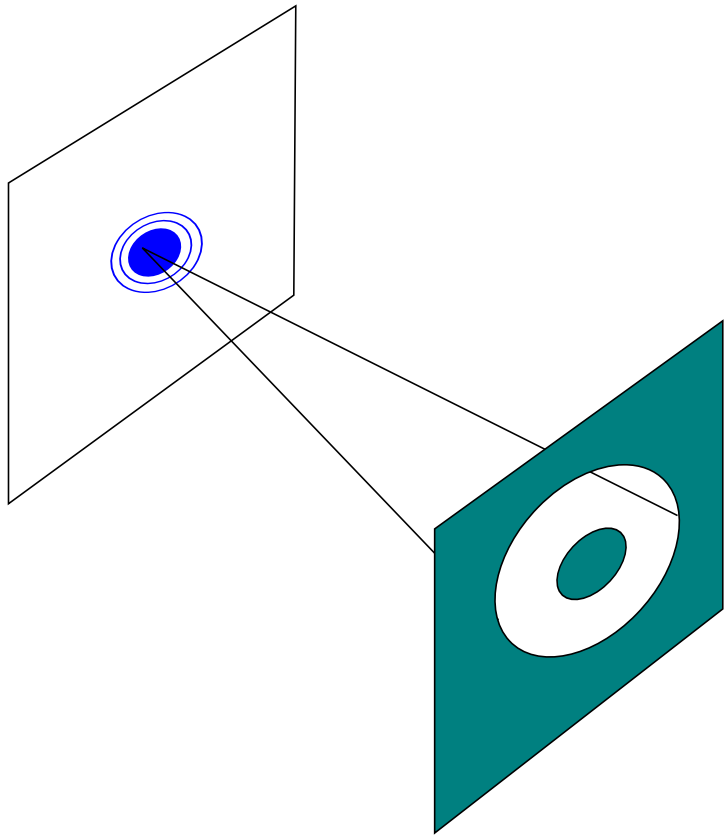

Figure 4.1: A circular aperture with a circular pupil mask producing an airy pattern in the image plane.

4.1 Aperture Masking Theory

Perhaps counter-intuitively the diffraction limited resolution of an optical system can be improved by placement of an obscuring mask over the aperture of the system. The basic setup is as seen in figure 4.1 in the simple case of a circular aperture and circular mask centred on the optical axis. The idea is to manipulate the input electric field in such a way as to counteract inevitable unwanted diffraction effects. It is easiest to understand the effect blocking off part of the aperture has within the formalism of Fourier optics and this will be used in this section to explain the situation.

It can be shown that the illumination of the image plane is simply the Fourier transform of the entrance pupil(Hecht 2nd 1987) for an object suitably far away to be considered in the Fraunhofer regime. We can set an arbitrary limit on this distance whereby the number of Fresnel zones encompassed by the aperture is much less than one.

$$\frac{D^2}{d \cdot \lambda} \ll 1 \tag{4.1}$$

Where D is the aperture of the optical system, d is the distance of the object and λ is the observing wavelength.

So given we meet this condition we can simply compute the output of any optical system via Fourier mathematics. Doing this for the elementary case of a point source at infinity and a circular aperture we get the familiar Airy pattern of a bright central spot and concentric diffraction rings. It is here that Fourier can give us insight into where these rings come from. The light from the point source will form plane waves at a suitably large distance so that the entrance pupil of the system is uniformly illuminated across the entire aperture. We can represent this mathematically with a top hat function Π, where $\Pi = 1$ for a radius of $\rho < D/2$ and $\Pi = 0$ for $\rho > D/2$, 1 within the radius $a = D/2$ and 0 outside. ρ being a radial vector from the centre of the aperture. We can look up the Fourier transform of Π and see it gives us a first order Bessel function, $\mathcal{J}_1$. To show this we begin with the result from studies of diffraction that the electric field E at a point P is

$$E = \frac{\epsilon_A e^{i(\omega t - \kappa R)}}{R} \iint\limits_{\text{Aperture}} e^{i\kappa(Yy - Zz)/R} \, \mathrm{d}\mathbf{S}, \tag{4.2}$$

where we introduce several new terms. E is the electric field disturbance at a point P on $[F]$ due to a source in the plane Σ. ϵ_A is the source strength per unit area, which we consider to be constant over the aperture, $\omega = 2\pi f$, t is some time of propagation and $\kappa = \frac{2\pi}{\lambda}$. f and λ are the frequency and wavelength of observation. R, Y, y, Z, z and $\mathrm{d}\mathbf{S}$ are all part of the geometry of the situation which can be seen in figure 4.2. The first term in equation 4.2 represents the strength of the disturbance at a distance R from the source and the integral representing a phase term taking into account the variation in phase caused by different optical paths from $\mathrm{d}\mathbf{S}$ to P.

Now as we have a circular symmetry it makes sense to transform from rectilinear x, y co-ordinates to cylindrical,

$$z = \rho \cos\phi, \;\; y = \rho \sin\phi$$
$$Z = q \cos\Phi, \;\; Y = q \sin\Phi,$$

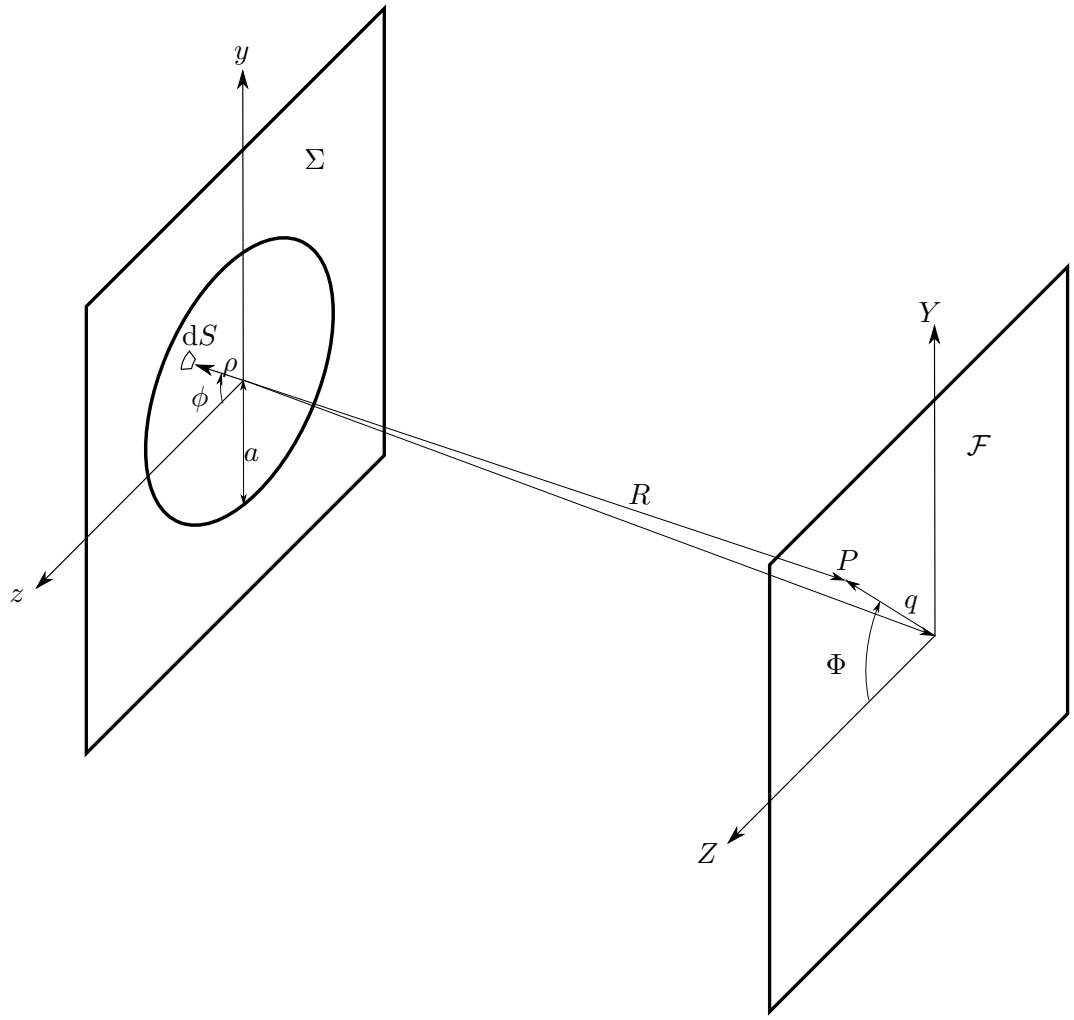

Figure 4.2: A circular aperture of radius a producing a diffraction pattern at a plane $\mathcal{F}$.

so that our source element becomes

$$\mathrm{d}\mathbf{S} = \rho \mathrm{d}\rho \mathrm{d}\phi.$$

so that 4.2 becomes

$$E = \frac{\epsilon_A e^{i(\omega t - \kappa R)}}{R} \int_{\rho=0}^{a} \int_{\phi=0}^{2\pi} e^{i(\kappa \rho q/R)\cos(\phi - \Phi)} \rho \mathrm{d}\rho \mathrm{d}\phi. \tag{4.3}$$

With the circular symmetry we can set $\Phi = 0$ and the integral over ϕ we recognise as the definition of a Bessel function,

$$\mathcal{J}_0(u) = \frac{1}{2\pi} \int_0^{2\pi} e^{iu\cos\nu}, \tag{4.4}$$

and 4.3 reduces to

$$E = \frac{\epsilon_A e^{i(\omega t - \kappa R)}}{R} 2\pi \int_0^a \mathcal{J}_0(\kappa \rho q/R)\rho \mathrm{d}\rho. \tag{4.5}$$

Using the recurrence relation of Bessel functions allows us to rewrite the integral in 4.5,

$$E(t) = \frac{\epsilon_A e^{i(\omega t - \kappa R)}}{R} 2\pi a^2 (R/\kappa a q)\mathcal{J}_1(\kappa a q/R). \tag{4.6}$$

The intensity is the focal plane is of course the real part of the field therefore we have the intensity distribution given by

$$\begin{aligned} I =& \frac{1}{2} E E^*, \\ I =& \left[\frac{\epsilon_A e^{i(\omega t - \kappa R)}}{R} 2\pi a^2 (R/\kappa a q)\mathcal{J}_1(\kappa a q/R) \right]^2, \qquad (4.7) \\ I =& \frac{2\epsilon_A^2 A^2}{R^2} \left[\frac{\mathcal{J}_1(\kappa a q/R)}{\kappa a q/R} \right]^2. \qquad (4.8) \end{aligned}$$

$\frac{2\epsilon_A^2 A^2}{R^2}$ is a constant for the source at a distance R and so we simplify this by setting it equal to $I(0)$. Identifying the angle q/R as the angle θ we finally arrive at

$$I(\theta) = I(0) \left[\frac{2\mathcal{J}_1(\kappa a \sin(\theta))}{\kappa a \sin(\theta)} \right]^2 \tag{4.9}$$

A slice through this bessel gives a squared sinc function as plotted in figure 4.3. And so the airy rings can be seen as the Fourier ringing due to the step in the Π function. This ringing is a natural consequence of either the Fourier mathematics used to describe the situation or more physically the effect of diffraction.

To see the effect of an obscuring disk as in figure 4.1 we remind ourselves of the Rayleigh criterion

$$\Theta = 1.22 \frac{\lambda}{D}, \tag{4.10}$$

But now we can see where this originates from. The arbitrary definition of equation 4.10 is the angular separation when the peak of one Airy disc meets the first zero of another. We can now see that this occurs where the term in the square brackets in 4.9 goes to zero,

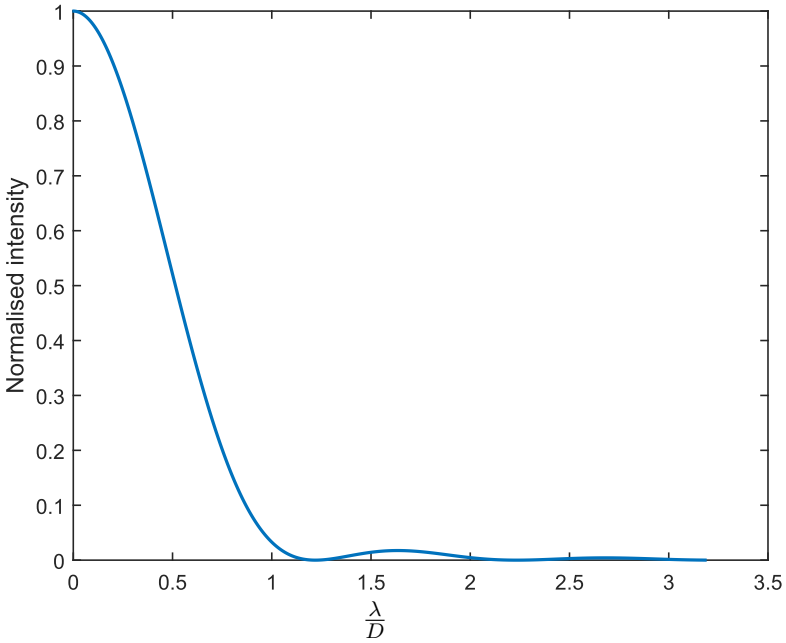

Figure 4.3: A slice through a bessel function shows the intensity rolling off from the maximum at the optical centre and reaching the first zero around $1.21\lambda/D$.

ie

$$\mathcal{J}_1\left(\kappa a \sin(\theta)\right) = 0, \tag{4.11}$$

and using the small angle approximation, $\sin(\theta) \approx \theta$, the argument in 4.11 rearranges to

$$\theta = x\frac{\kappa}{a} \tag{4.12}$$

$$\theta = \frac{x}{\pi}\frac{\lambda}{2a} \tag{4.13}$$

$$\theta = \frac{x}{\pi}\frac{\lambda}{D}, \tag{4.14}$$

where $\mathcal{J}_1(x) = 0$ for $x = 3.83, 7.016, \ldots$. And we find $3.83/\pi = 1.22$ and we recover (1.1).

We can now calculate the angular resolution for a circular aperture with a central circular mask of some diameter $r = \epsilon D$ where $0 \leq \epsilon \leq 1$ and D is the diameter of the full aperture. Then from equation (4.3) with the radial integral only over the domain from $\epsilon a \leq \rho \leq a$, where $a = D/2$, and following the same steps to (4.9) we get

$$I(\theta) = \frac{I(0)}{(1-\epsilon^2)^2}\left[\frac{2\mathcal{J}_1(\kappa a \sin(\theta))}{\kappa a \sin(\theta)} - \epsilon^2\frac{2\mathcal{J}_1(\kappa\epsilon a \sin(\theta))}{\kappa\epsilon a \sin(\theta)}\right]^2. \tag{4.15}$$

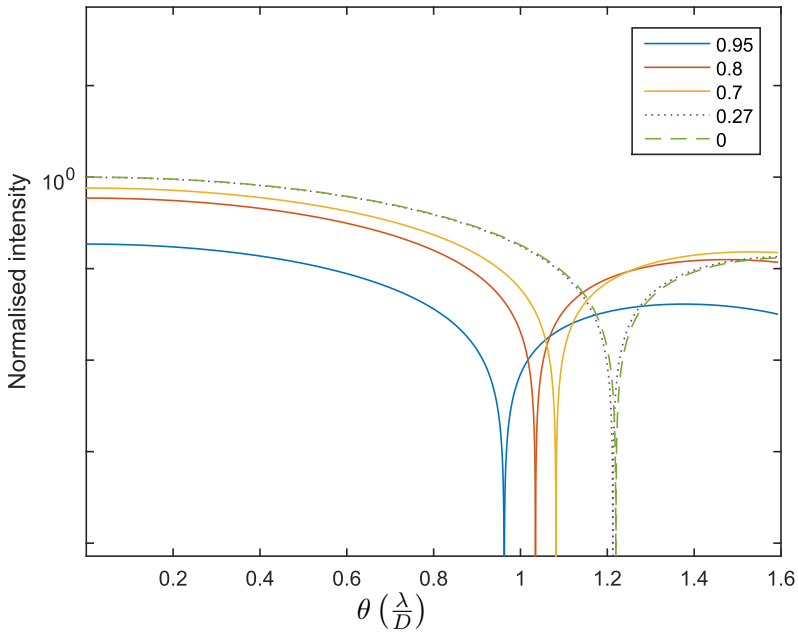

Figure 4.4: This figure represents the intensity profile at focus for an circularly obscured aperture, radius a. Each curve represents a different value of ϵ where the mask is of radius ϵa. We see that the first zero gets progressively closer to zero as the mask increase in size and that the central intensity reduces as the aperture becomes more obscured. Included is the profile for a mask of $\epsilon = 0.27$ that correspond to the inner radius of the mask leading to the result in figure 4.12.

Searching for the roots of of equation (4.15) we find that

$$\Theta = 1.00\frac{\lambda}{D}, \tag{4.16}$$

which tells us we can resolve features of slightly smaller size when using a mask according to our Rayleigh criterion. Plotting the intensity I against angle θ in units of λ/D for varying epsilon gives us figure 4.4. Here the intensity is plotted logarithmically to show clearly where it comes to zero for 5 different values of epsilon; $\epsilon = 0.95, 0.80, 0.70., 0.50$ and 0.00. We can see how increasing the size of the obstruction, increasing ϵ, reduces θ and thus improves our resolution. It is important to note the denominator in equation (4.15) contains a $1 - \epsilon$ term and so the intensity is reduced as might be expected.

Further results for more complicated pupil plane distribution can be calculated and these are shown in figure 4.5. We can see we have the power to favourably alter the image plane through manipulation of the entrance pupil, indeed figure 4.6 shows the well known effect of star "points" resulting from diffraction off the secondary support of a Newtonian telescope.

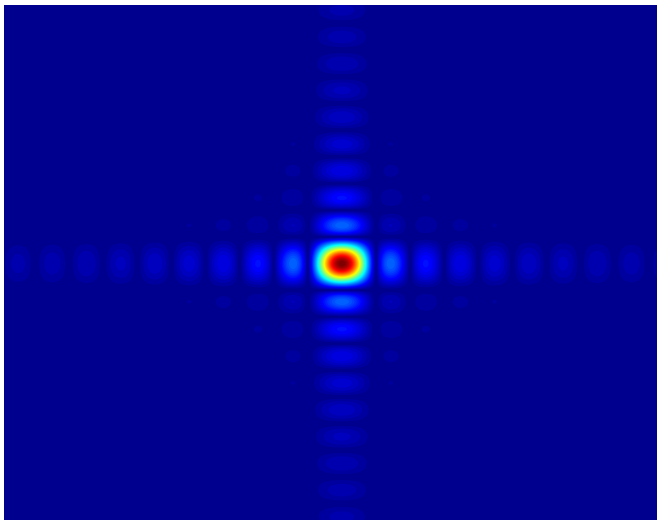

Figure 4.5: A square aperture shows diffraction ringing perpendicular to the edge encountered.

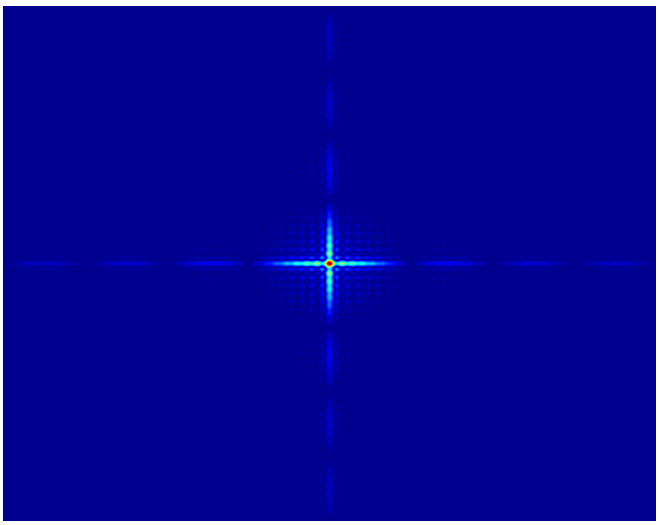

Figure 4.6: A cross aperture shows the familiar star spikes of diffraction from a spider in the telescope aperture.

4.2 Apodised Pupil Masking

The purpose of my work in this field was to enhance the imaging of exoplanets via a pupil mask that reduced the airy rings of objects in the field, bringing down the size of the Airy disc as seen in equation 4.16. The concept being inspired by the work of Kasdin(Kasdin et al. 2003) et al. The basic design being a petalled "star-shade", these petals are aimed at reducing the ringing via apodisation of the disc in the pupil.

Apodisation is manipulation of the pupil to remove the step functions at the edges of the aperture and replace them with Gaussian profile(Tschunko 1983). Again this is explained by Fourier, as the transform of a Gaussian profile is another Gaussian and as such there is no ringing, just a bright central spot. The aperture function in 4.2 would be a Gaussian for this technique. Unfortunately a Gaussian profile is impossible to recreate in the lab as the wings continue to infinity and therefore any physically realisable aperture has to truncate the Gaussian. Allied with this is the mechanical difficulty of reducing the throughput of the aperture with an exact analytical profile.

These two problems have stunted the exploitation of the field so far and most work in this area is theoretical outside of the work of Kasdin. To refine and reduce the problem of generating a mask a measure of fitness was devised. The measure taken was the reduction in height of the sidelobes of the airy pattern produced assuming a circular aperture enclosing the mask. So a test pattern could be produced and the output of any mask could be compared to this original unobscured aperture. Each mask generated could be assessed mathematically and a optimum could be found.

This approach limits our confidence in any optimum found due to relatively few masks, $\mathcal{O}(100)$ being assessed. This limit was due to available processing power.

4.3 Apodised petal mask design code

The higher level description of the code can be seen in figure XXX and in some detail in appendix C. The first part sets several variable constituents of the design such as petal height, size of central obscuration and number of petals. These define the parameter space to be searched via an iterative process outputting one dimensional slices of the calculated image plane illuminated. Left to run for some period of time several designs could be evaluated and fine tuning could be directed at the masks showing most promise of reducing the intensity of the first airy ring. Typically a run of 20 masks could be

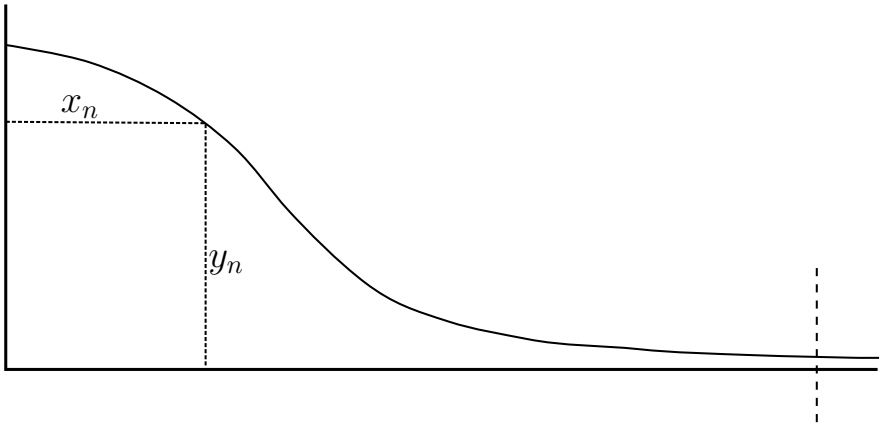

Figure 4.7: A scaled Gaussian curve is digistised by taking points along the curve. The number of points, M depends on the required resolution, ie how squared off can the curve be allowed to get.

completed within 12 hours.

4.3.1 The code in detail

To calculate the petal curve the code draws a curve in (x, y) to represent half a Guassian from the peak to some value arbitrarily close to zero. This curve is digitised by taking y-values on this curve at a set number of points along the x-axis. The logic being that for each petal on the outside of the central disc the curve will be tangentially offset from some radial line as shown in figures 4.7 and 4.8. We now have a petal shape with a Gaussian edge profile mapped in (x, y), however the circular symmetry strongly suggest converting to polar co-ordinates (r, θ), where

$$r = \sqrt{x^2 + y^2} \tag{4.17}$$

$$\theta = \tan^{-1} \frac{y}{x}. \tag{4.18}$$

So armed with the polar co-ordinates of one petal we can now readily repeat this petal around the circumference of the central mask as many times as our design requires. The number of points defining the curve is set as M, so that the resolution can be varied depending on computing power or to simulate machining errors. r will remain constant

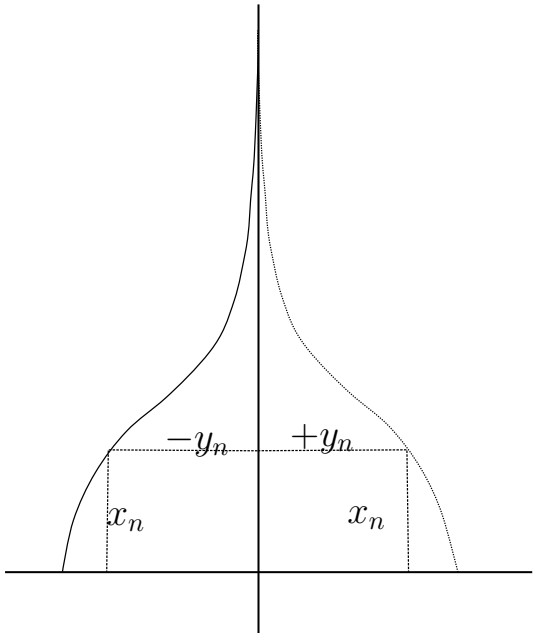

Figure 4.8: The digitised curve is rotated and reflected to create a standard *petal*. This petal has a user definable resolution so that the average distance from an analytic curve can be controlled for. Manufacturing errors can be simulated by varying this "squaring off."

for each point on the petal and θ will change by[1]

$$\theta \bigcup n\frac{2\pi}{N}, \text{for } n = 1 \text{ to } N-1 \tag{4.19}$$

where n is the n^{th} petal and N is the the number of petals in the design and we approach the situation in figure 4.9

The petals were coded so that the basic structure could be varied. The principle parameters for variation other than the total number of petals, N, were the radius of the inner circular mask, r, and the height of the petal from the edge of the central mask, h. This is shown pictorially in figure 4.10. The curve of the petal, though chosen here to be Gaussian, is user definable and can be modified for optimisation. As this work used a relatively inefficient evaluation method the curve was left as Gaussian.

An example of the output of the code is shown in figure 4.11 To evaluate this mask it was assumed that the mask sat directly in the entrance pupil of a telescope and that it was perfectly binary. So that the area defined within the petals and central disc was perfectly opaque and the area outside the mask but inside the aperture was perfectly

[1]The actual coding is slightly more complex: $\theta_n = \frac{\pi}{2} + n\frac{2\pi}{N} - \frac{\pi}{N} + \frac{m2\pi}{MN}$, with a double index over m and n.

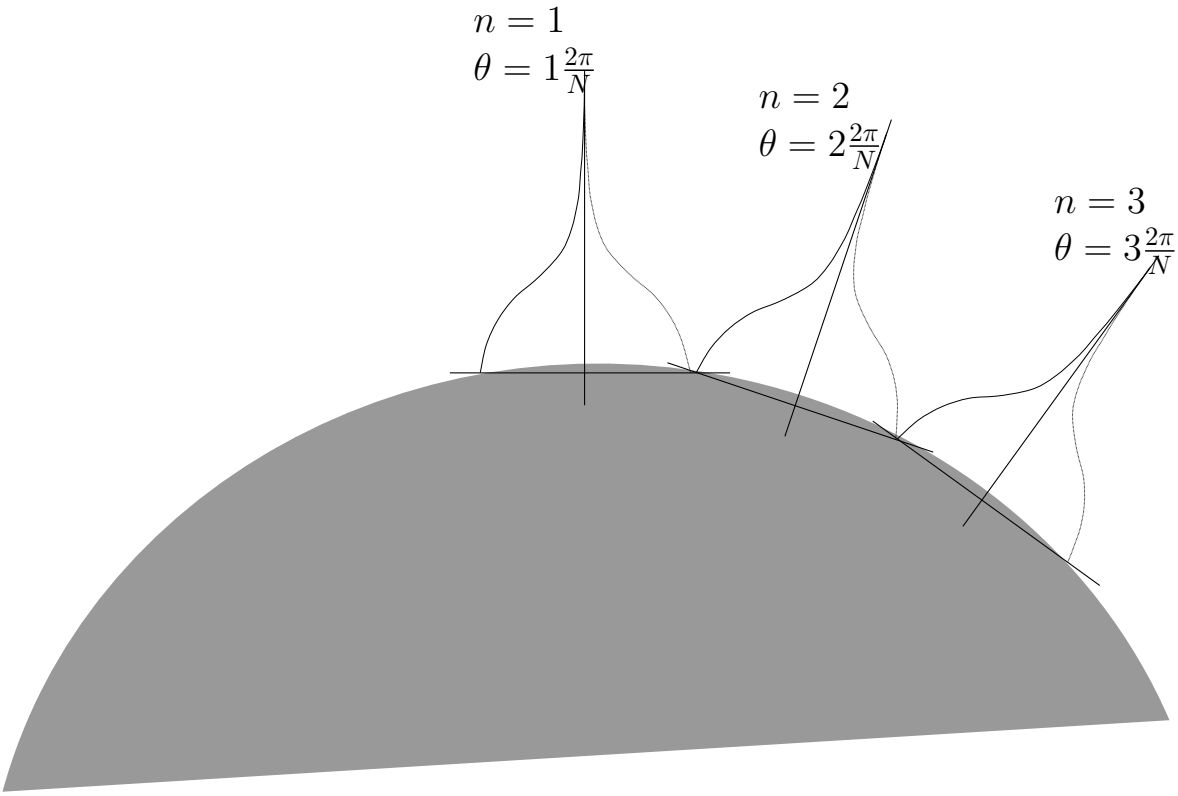

Figure 4.9: The petals once converted to polar co-ordinates could be scaled and applied to the edge of the disc N times as per the desired number.

transmitting. Using code from the online Matlab database the mask was converted from a vector detailing the r, θ co-ordinates to an $x - y$ transmission map. This map represented transmission with the value 1 and obscuration with 0. In accord with the theory in section 4.1 the electric field in the focal plane is the Fourier transform of this map. To do this the two dimensional fast Fourier transform function within Matlab was used, *fft2*. The output of this places the zero frequency in the upper left corner, to better represent the physical situation at the detector *fftshift* was used to move this to the centre of an output matrix. When the piecewise-square of this matrix is taken we end with the intensity distribution at the output of the system. The assessment of this output is covered in the following sub-section.

4.3.2 The evaluation of the code output

To evaluate the code it is useful to identify a metric against which a particular mask configuration can be assessed.

In table 3.4 and making use of the theory in 4.2 we can understand that to probe the angular scales required we might get there using an apodised pupil mask. So that

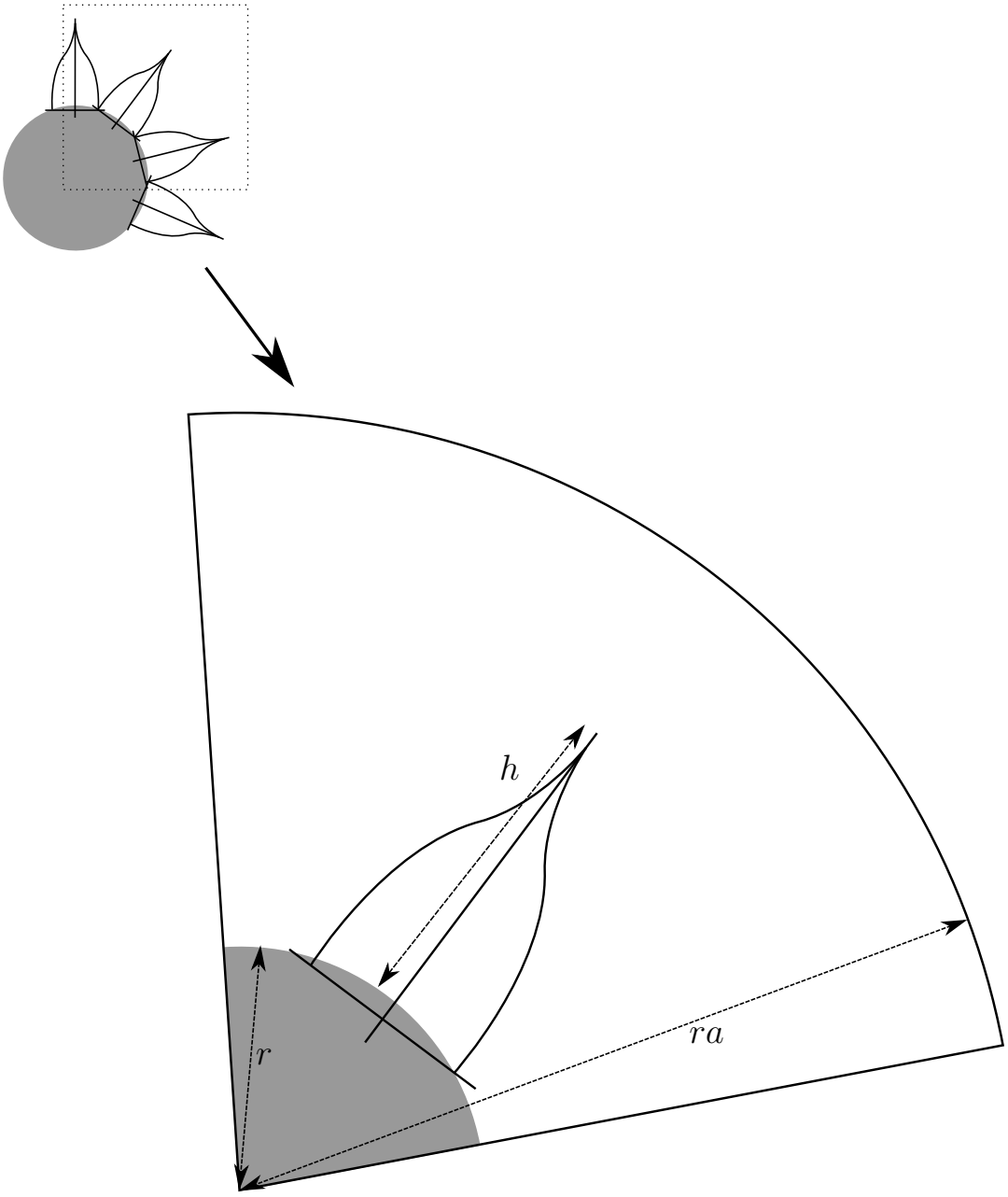

Figure 4.10: A few of the apodising petals are shown top left. Enlarging a portion of the mask we can see the variable parameters within the code. The radius of the central mask r, the additional height above this of the petal, h, and the size of the aperture that we are masking, ra. As well as these parameters the function of the curve can be altered to the desired form and the total number of petals can chosen.

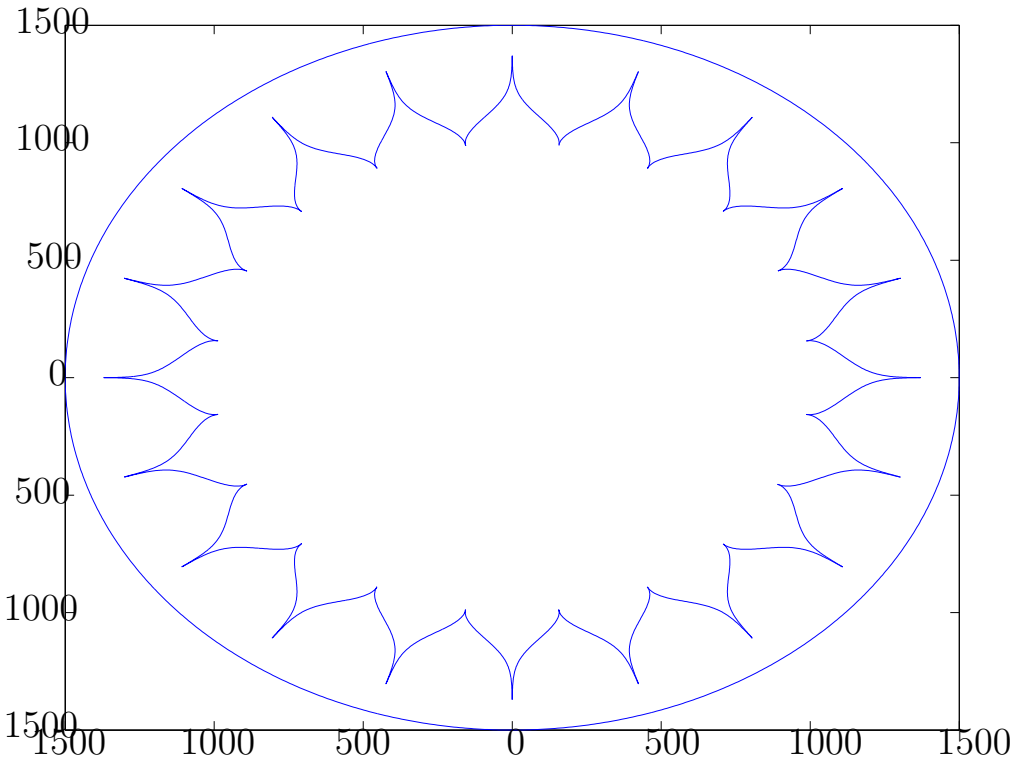

Figure 4.11: The output of the petal code is shown here. Parameters as defined in the text as follows: $N = 20$, $r = 400$, $ra = 1500$, $r = 1000$

the primary goal of using the mask is to reach an angular resolution of our instrument sufficient to resolves distances that allow for exoplanetary investigation. This then leads us to see that the distance in the focal plane to the first minima of an unresolved star should be reduced to an equivalent angular size of around 10 to 50 milli-arcseconds.

And so taking the output as shown in figure 4.12 we take a slice through the centre of the image to get a one dimensional vector. The plot of this we might expect to look similar, if not identical, to that shown in figure 4.3. With the mask present however we find that the sinc function has been modified with the first zero closer to the centre of the plot. Figure 4.12 shows an overlay of the modified sinc function against the sinc function given by an aperture of the same radius as the outer aperture in 4.11, ie $= ra$ with a simple circular mask of radius r installed. The x-axis is marked in arbitrary units centred on 1000. This allows us to properly evaluate the effect of the apodising petals, rather than simply re-iterate that a masked aperture brings the first zero closer to the photocentre of the diffraction pattern.

We see that with a central circular mask of radius 400 with a petal height of 640 does bring in the first zero by a few points but there is still significant ringing, though not worse

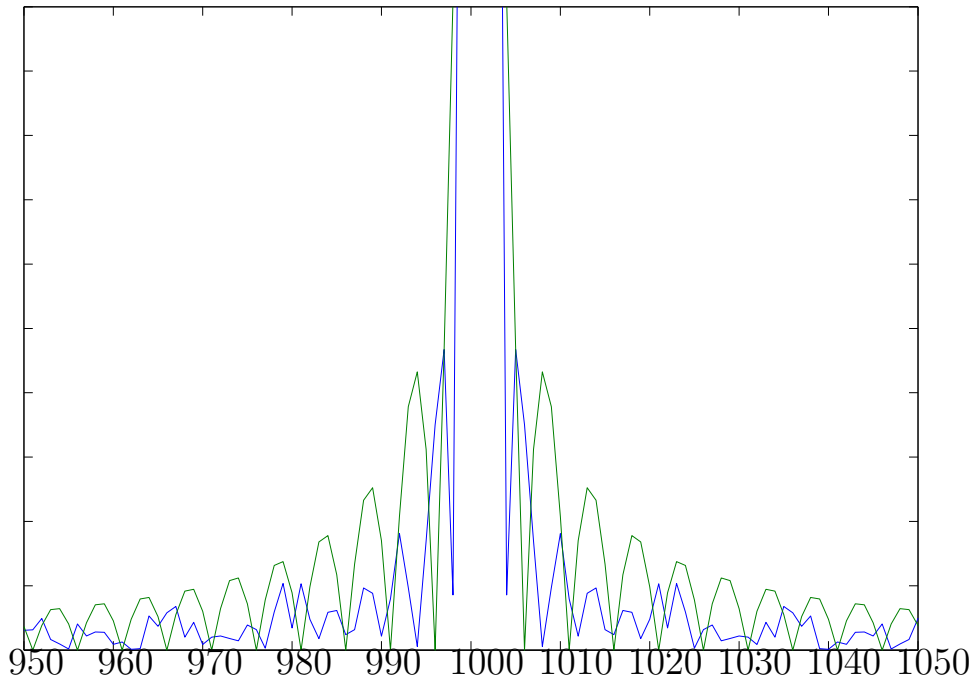

Figure 4.12: A slice through the calculated focal plane gives a trace of the intensity. The Blue curve represents the diffraction pattern with a petal mask installed. The Green curve represents the intensity with a simple circular mask installed. Parameters as defined in the text as follows: $N = 15$, $r = 400$, $ra = 1500$, $h = 640$. Note the figure is zoomed the area of interest to show where the function approaches zero. Both curves are normalised.

than with a simple circular mask. And this ringing drops off significantly quicker. Which suggests we may get some imaging improvement with use of the mask. The first zero may be seen to be at a distance approximately 3/5 of the first zero for the plain circular mask. This plain mask has a resolution limit at just under $1.21\left(\frac{\lambda}{D}\right)$, so that very roughly we may say the apodised mask has a resolution at

$$\begin{aligned}
\theta_{apodised} &\approx \frac{3}{5} \times \theta_{circular} \\
\theta_{apodised} &\approx \frac{3}{5} \times 1.2 \left(\frac{\lambda}{D}\right) \\
\theta_{apodised} &\approx 0.72 \left(\frac{\lambda}{D}\right), \qquad (4.20)
\end{aligned}$$

which shows a modest improvement. Whether this improvement is significant in an imaging depends strongly on how the diffraction wings affect the overall picture. This is not covered here and awaits further work.

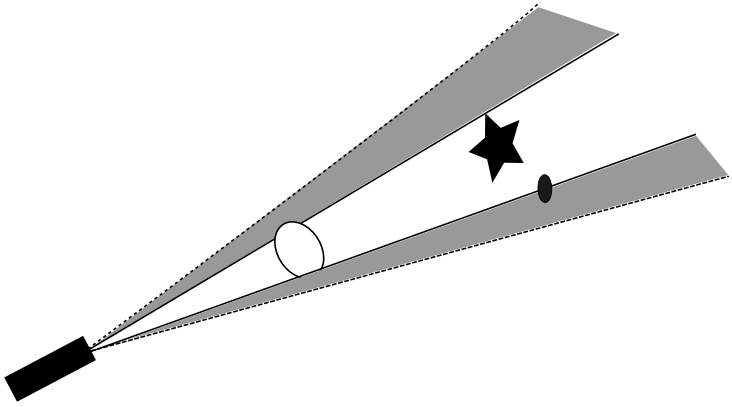

Figure 4.13: A star occulter flies in front of the telescope in the bottom left corner and is positioned so that it shadows the star without blocking the direct path to an orbiting planet. The region encompassed by the total field of view, in one dimension, with the shadowed region subtracted is shown by the grey regions. In practice this zone creates an annulus around the occulter so that a planet's full orbit may be observed.

It is important to note that the throughput of the system is strongly reduced, to approximately 40% with these parameters. This is less of a problem when viewing bright sources and there are spare photons. Also this reduction in throughput also reduces problems associated with scattered light.

4.4 Apodised petal mask applications

The original mask design has been designed to be a star occulter free flying in front of the James Webb Space Telescope at a significant distance while the design described in this chapter is to be attached at the pupil of a telescope. The star occulter is briefly described before the outline of a space telescope with apodised pupil mask is given.

4.4.1 Star occulter

A star occulter differs from a pupil mask in that it is placed to geometrically shadow the telescope aperture of the star light but not of a planet in orbit. This requires precise placing of the mask so that it's angular size is larger than the host star but less than the angular size the planets orbital distance as shown figuratively in 4.13 The positioning of the occulter can be calculated simply from the triangle formed by the linear size of the host star and it's distance, we have seen in table 3.4 that a typical angular size of a host star may be $0.250\text{milli} - \text{arcseconds}$ and 1AU is $50\text{milli} - \text{arcseconds}$. Therefore we require our occulter to have an angular size, θ, between these values. It is easier to fly the occulter

at nearer distances than further so that we would like to maximise the angular size and we choose a size of 25milli − arcseconds. A further constraint is that the shade must have a diameter, D, larger than the telescope aperture, this gives us a base for our long thin triangle and in this case we may choose 5 metres. This ensures a 3 metre aperture is well shadowed. Thus the distance, r, to our occulter is given by

$$r = \frac{D}{\theta} \tag{4.21}$$

$$r = \frac{5}{25 \times 10^{-3}}$$

$$r = 4.13 \times 10^{7}\text{m}$$

$$r = 41300\text{km}. \tag{4.22}$$

Obviously 40000 kilometres sounds like a long way away and indeed this precludes a low Earth orbit but in more distance orbits this is not an insurmountable problem. It is still a problem, not least with station keeping and repositioning.

There are benefits of the system over a pupil mask acting as a coronagraph, as the star light that enters the telescope's entrance pupil is much lower and the major problem of scattered light is very much reduced. Optimised schemes such as though planned for operation with a JWST-like telescope such as the New Worlds Observatory/Terrestrial Planet Finder - Occulter(Puig et al. 2010) reckon on a suppression of the incident starlight at focus by a factor of around 10^{-10}.

4.4.2 Apodised pupil mask

In this work the petal mask shape is to be used as a pupil plane mask whereby the petal shapes act to apodise the mask. As already described several times the mask is designed to lie in the pupil plane of a telescope. As such in exo-planetary work it would operate much like a coronagraph. For this mission concept a space telescope operating in the far infrared near to $\lambda = 250\mu$m. The aperture d would be three metres in size and use a cryogenic detector system as seen in the Herschel Space Telescope. With such a setup the unobstructed, ie no mask, diffraction limited angular resolution θ_{DL} of the system would

be expected to be

$$\begin{aligned}\theta_{\text{Herschel\_DL}} &= 1.22 \times \frac{\lambda}{d} \\ &= 1.22 \times \frac{250\mu\text{m}}{3\text{m}} \\ &= 100.0 \times 10^{-6} \text{ radians} \\ \theta_{DL} &= 21.0''. \end{aligned} \tag{4.23}$$

With the mask seen in figure 4.11 and using the calculations in (4.20) we might expect our resolution to increase to approximately 3/5 Herschel_DL,

$$\begin{aligned}\theta_{\text{masked}} &= 3/5\theta_{\text{Herschel\_DL}} \\ \theta_{\text{masked}} &= 12.6''. \end{aligned} \tag{4.24}$$

Again it needs to be noted that this is just a raw improvement on one metric, whether this is seen as an improvement overall depends on the scientific objective of the instrument.

For the mission of observing exoplanets we are still a long way from investigating the search space as defined in table 3.4. As such this work can be seen as of limited possible exploitation in regards to direct viewing of exoplanets. Another application however could be the imaging of proto-planetary discs.

These discs are large structures of gas and dust that are orbiting their host star in a thin disc, at a stage just prior to the formation of planetary solar systems. They can stretch as far as 50 AU in radius and therefore are much larger targets. If we wanted to place 5 resolution elements across the image then we might hope to resolve such detail out to a range of 1.5pc. This is a distance would still limit the search space and considerations of flux throughput are again important so that we may consider only bright sources.

4.5 Shaped Pupil Design

An alternative design that was then looked at is the Spergel mask(Kasdin et al. 2005). This is an unapodised pupil mask that seeks to alter the contrast ratio between two objects in the field of view. This is done by dispersing the light to particular zones of the image plane depending on the distance from the optical axis. Those objects lying on or near to the optical axis have their light concentrated in two symmetric zones about the image

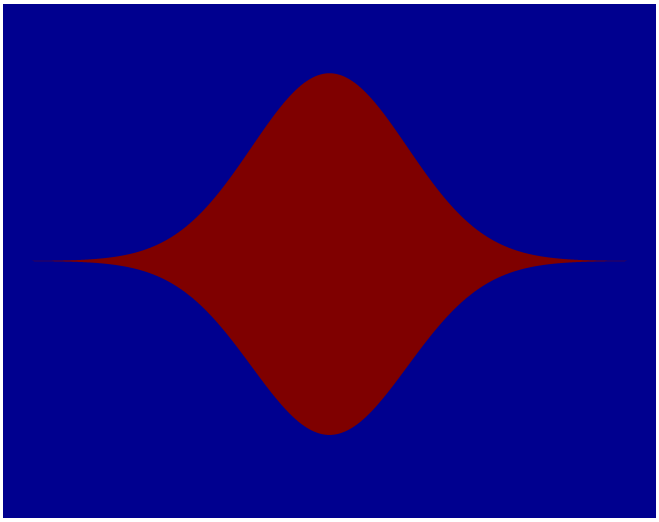

Figure 4.14: Pupil plane illumination of a Spergel mask. Red denotes transparent(= 1) and blue opaque(= 0).

plane while off axis objects are imaged largely unaffected due to the contrast ratio between the star and object of study. As seen in table 3.4, section 3.2 the this ratio is substantial and carries the threat of swamping off axis sources as the diffraction wings get blown up.

The purpose of this is to aid the viewing, direct imaging of exoplanets or spectrographic analysis. As seen in section 3 the contrast ratio between a planet and its host star can be of the order 10^7 at $\lambda = 20\mu$m. This is typically of a dynamic range beyond current detectors reach. And so with our Spergel mask we hope to reduce this contrast in particular "detection zones" in the image plane. The resulting image plane illumination can be explained by reference to the Fourier optics as described above. Figures 4.14 and 4.15 as generated in MATLAB show how the flux is diverted to a cross shape centred on the centre of the image. Light from an off axis source is affected in the same way, the point spread function (PSF) has the same shape, and we would see this light brought to focus as expected. However while the off-axis object sits in a discovery zone the contrast ratio is improved. The measured improvement in contrast ratio is of the order of 10^2 and we might expect to be able to discern the ratio of 10^4 with modern technology. Equation (3.10) and figure 3.10 tell us we face a ratio of the order 10^7. Therefore the design fails to meet this requirement. The main failure to achieve the required reduction in flux ratio comes from the digitisation of the curve in the Matlab code. To carry out a finite Fourier transform the curve needs to be squared off. The resolution of the Fourier transform is

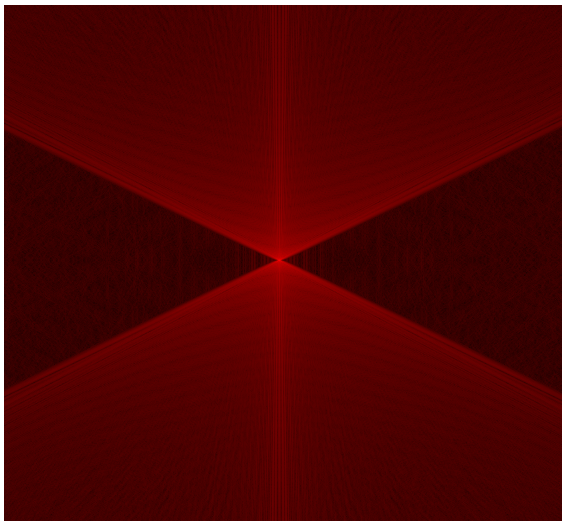

Figure 4.15: Image plane illumation. Red denotes illuminated portion of the plane. Note the two detection zones, left and right of the image centre.

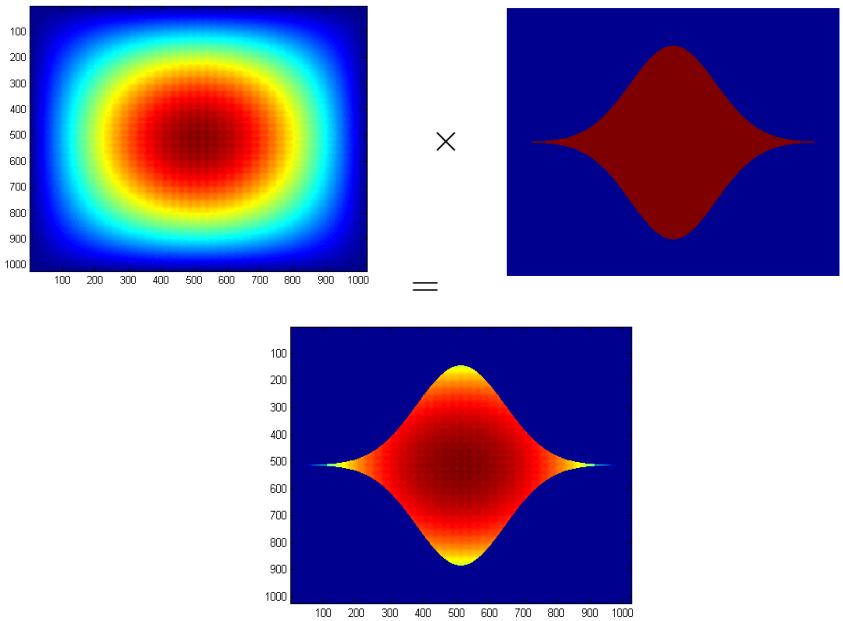

Figure 4.16: The figure shows top right the electric field as produced by a bright central object and two faint objects (brightness ratio 100:1). One positioned due South and one due East. Top right is the Spergel pupil. At the bottom is the product of the two, the physical electric field that will propagate through our system.

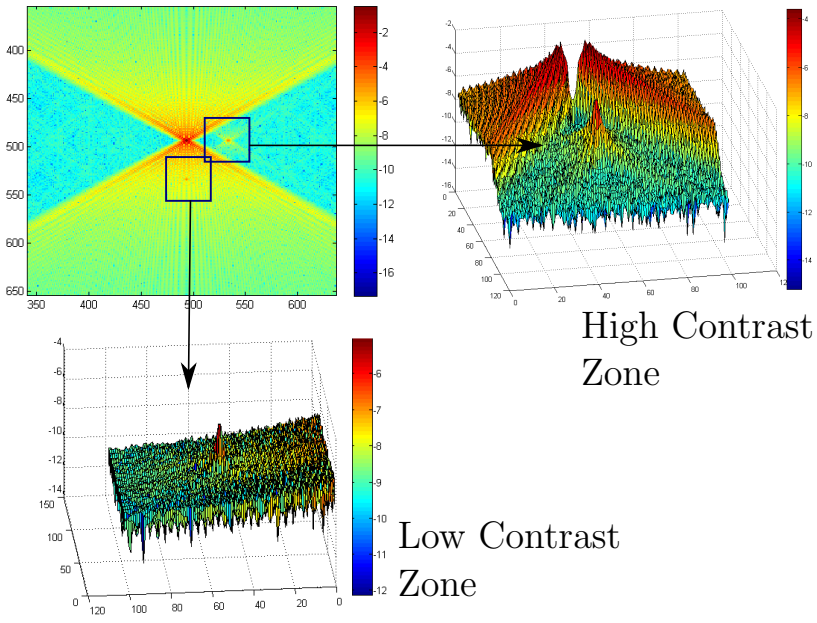

Figure 4.17: Top left shows the Fourier transform of the lower image in figure 4.16. The yellow and red areas show where the light from the "star" has been diffracted too and the blue areas where the light has been diffracted from. The object to the South is in an area of low contrast as shown in the surface plot below, object to the East is in a high contrast region.

then limited by the processing power available. Reflecting this, the concept has not been employed in the field over concerns about the accuracy that the mask could be constructed.

Any deviation from an analytic curve and the desired contrast ratio reduction is significantly reduced. As seen in figure 4.15 the contrast reduction in illumination within the detection zone is of $\mathcal{O}(10)$, well short of the analytic capability. The reason for this is the "graininess" as seen in figure 4.14. The curve is squared off within the code and assuming the mask is one metre wide at its widest point vertically then the pixel size is equivalent to 200μm, which is a considerable engineering challenge already. Secondly, the flux throughput of the system is necessarily reduced due to the opaque nature of the mask. A further problem with the system is that not all of the image plane benefits from a reduced contrast and as such you have a reduced detection efficiency over other methods. If the on axis light is constrained to 10% of the image plane then your time spent looking for exoplanets is also reduced to 10%. As such this method presents some very tempting benefits. Further iterations are possible where the shape is repeated vertically, each iteration will increase the search areas(Kasdin et al. 2005).

4.5.1 A space based aperture masked spectrograph

Unmentioned so far in this chapter is the investigation of the atmospheric composition of exoplanets. In the search for extra terrestrial life finding hospitable planets is only the beginning. The primary method of identifying biological processes on these planets seems to be looking for bio-marker molecules in their atmospheres. This is a large and broad subject that isn't discussed further here other than to list several of these molecules that we might be of interest in table 4.1.

Bio Marker	Strong lines present at: μms
CO	2.3, 4.6
CO_2	2.8, 4.3, 15
CH_4	3.3, 8.0
NH_3	3.0, 6.1, 10.5

Table 4.1: Table of significant bio-markers that we may use as targets for spectroscopic analysis of exoplanetary atmospheres.

What can be described are instruments that might help in collection and analysis of light from atmospheres such that spectroscopy can be carried out. The problem that limits current investigation is that the light received at Earth is with current instruments is not spatially resolved. With transiting exoplanets we can deduce a planetary atmosphere via spectrographic substitution whereby the light from the host star is subtracted from host star and planet. This light is obviously collected at different parts of the planet's orbit such that it is obscured at some point. This is still a messy method as the light from atmosphere, includes reflected as well emitted radiation. What can help through spatially resolving the source is the light from the planet can be sent to a spectrograph separately. As discussed at length above the ability is very challenging and it is proposed that this could be achieved through pupil masking.

The flux ratio $\mathcal{R}$ between the host star and an exoplanet at 1AU is calculated above, equation (3.10), and so we need to separate the star from the planet and reduce the diffracted light. This diffracted light needs suppression by a factor of at least $\mathcal{R}_{\text{wing}}$, the ratio of star light to the planetary light in the wing of the of the image. Considering the star image to be an airy disc then the further the separation of the planet from the star the lower this ratio becomes, given some modulation. However we can shape the pupil of our telescope such that the result we obtain in the image plane is more helpful.

The shape of pupil used in this work was devised by Kasdin et al.(Kasdin et al. 2005).

This pattern distributes the light into a star shaped pattern as seen in 4.15. This immediately helps with $\mathcal{R}_{\text{wing}}$ by redistributing the light from some "detection" zone to elsewhere in the image. Analytically this method can reduce the flux ratio in the detection zones by a factor of 10^6 in the NIR. With this reduction we can apply some version of an integrated field unit whereby the light from different points in the image plane can be directed to a spectrograph separately[2]. This allows for a direct measure of the spectra of exoplanetary atmospheres.

With such a mask placed in front of a 12 metre telescope we could expect to resolve the habitable zone around host stars out to a distance of around 6 parsecs. This distance would limit the ability to conduct wide surveys but allow for a high accuracy of spectrographic measurement.

[2]See section **??** in chapter **??** for one such way to do this.

This page was intentionally left blank

Chapter 5

A labtop Intensity Interferometer

Robert Hanbury Brown did in 1950(Brown and Twiss 1957)(Brown and Twiss 1958) what all good physicists do and abstracted the ideas associated with practical interferometry, that is the recombination of light *in phase*, and reconsidered which physical phenomena had to be recombined. In previous discussions about the procedure of interferometry we have made reference to the phase of the electric field being recombined and noted that this leads to a coherence time Δt and coherence length Δl(Haniff 2007) of around 1μm. It is this tough engineering constraint that Hanbury Brown thought to overcome, and he did this by using the correlation of the phase of the intensity φ_I in place of the phase of the amplitude φ_A.

It is by no means obvious that there should be a correlation measurable within φ_I across two apertures but with mathematical assistance provided by Richard Twiss a coherent theory was developed and proven in the lab over the next decade(Brown and Twiss 1958). Some counter experiments were attempted to contradict this theory but they came up short and the *Hanbury Brown and Twiss effect* is accepted science today[1]. As well as being used in astronomy it has also found applications in the field of particle physics. Photons are bosons and particle physics of course has its own interest in that class of particle, the quantum physics is directly transferable.

In this chapter a brief update of the theory of interferometry with respect to measurements of φ_I instead of φ_A is given. Then there is a description of laboratory work to

[1]Indeed Roy Glauber's work to determine the concept from a quantum physical perspective earned him the Nobel Prize in 2005(Glauber 2006)

build an updated Intensity Interferometer as an alternative instrument architecture to the orthodox amplitude based interferometers.

5.1 The theoretical underpinnings of the Intensity Interferometer

To explain the use of the phenomena in astronomy it is not necessary to invoke quantum mechanics, we can use the classical wave mechanics as described in chapter 2 with a small amendment to the interfered quantity.

An electromagnetic wave is typically described by a magnitude and an oscillating phase in such a way $E\varphi^{i\theta}$. However a thermal source will have a non constant magnitude and as such we need to consider E as complex as well with a separate phase, denoted as $E\varphi_I$ or $\tilde{E}$. It can be observed that φ_I oscillates on a much slower timescale than φ_A, by a factor of 10^4. Which leads directly to a coherence length increased by the same factor.

This increase in coherence length is of significant benefit to the construction of an interferometer. As can be shown(Haniff 2007)(Dravins et al. 2012) the total induced path length error must be less than half the coherence length

$$\Delta L < \Delta l/2.$$

The increase in Δl by a factor of 10^4 means the construction of the instrument is constrained to being accurate to 30cm against a limit of 1μm. [2] Building to these limitations is obviously an easier task and this de-restricting of build quality lets us be much more optimistic about the viability of a full size instrument.

To understand how an Intensity Interferometer works I use the explanation given in the introduction to "The Intensity Interferometer" written by Robert Hanbury Brown(Hanbury Brown 1974), further explanation can be found in (LeBohec et al. 2008). We refer back to figure 2.2 where two points radiate independently and interfere at two apertures but we replace the apertures with detectors. Here we move the discussion further by considering the current created in the detectors by the incident radiation. A simple consideration about the behaviour of the detector would suggest the current produced is proportional to the intensity of the light I. Therefore the current produced, i_A at detector A, i_B at detector

[2] A path length error of 30cm would lead to a reduction in the interferometric visibility of 50% when observing in the optical.

B, will be the sum of the electric field disturbances created by σ_1 and σ_2:

$$i_A = K_A \left[E_1 \sin(\omega_1 t + \phi_1) + E_2 \sin(\omega_2 t + \phi_2)\right]^2 \tag{5.1}$$

$$i_B = K_A \left[E_1 \sin(\omega_1 (t + d_1/c) + \phi_1) + E_2 \sin(\omega_2 (t + d_2/c) + \phi_2)\right]^2 . \tag{5.2}$$

The electric field at B showing an extra phase term to represent the longer distance the light has travelled to B than the path to A. And we use K to represent the efficiency of the detectors in converting photons to electrons.

And so multiplying out the squared brackets we find that both i_A and i_B have a similar form of four terms;

$$\begin{aligned} i_A =& \frac{1}{2}\{(E_1^2 + E_2^2) - [E_1^2 \cos 2(\omega_1 t + \phi_1) + E_2^2 \cos 2(\omega_2 t + \phi_2)] \\ & - 2E_1 E_2 \cos((\omega_1 + \omega_2)t + (\phi_1 + \phi_2)) \\ & + 2E_1 E_2 \cos((\omega_1 - \omega_2)t + (\phi_1 - \phi_2))\}, \end{aligned} \tag{5.3}$$

$$\begin{aligned} i_B =& \frac{1}{2}\{(E_1^2 + E_2^2) - [E_1^2 \cos 2(\omega_1 (t + d_1/c) + \phi_1) + E_2^2 \cos 2(\omega_2 (t + d_2/c) + \phi_2)] \\ & - 2E_1 E_2 \cos((\omega_1 + \omega_2)t + \omega_1 d_1/c + \omega_2 d_2/c + (\phi_1 + \phi_2)) \\ & + 2E_1 E_2 \cos((\omega_1 - \omega_2)t + + \omega_1 d_1/c - \omega_2 d_2/c (\phi_1 - \phi_2))\}. \end{aligned} \tag{5.4}$$

The first is an expected DC offset representing the basic response of the semiconductor. The second term is more subtle and represents the second harmonic of the combined radiation fields. This term is filtered out from the signal and does not reach the multiplier. This electronic chain is explained next in section 5.3. The third term again can be filtered by our electronics.

The fourth term however is of use to us. This term causes a fluctuation in the intensity of the light that we expect to be correlated as it is the sum of the electric fields, which we can see by multiplying the fourth term of equations (5.3) and(5.4), letting $c(d)$ stand for the correlation.

$$c(d) = K_A K_B E_1^2 E_2^2 \cos[(\omega/c)(d_1 - d_2)]. \tag{5.5}$$

As the phase offset $(d_1 - d_2)$ increases in magnitude with angle θ the correlation will drop to zero. It can then be seen by direct analogy to 2.2 and (2.4) that our Intensity Interfer-

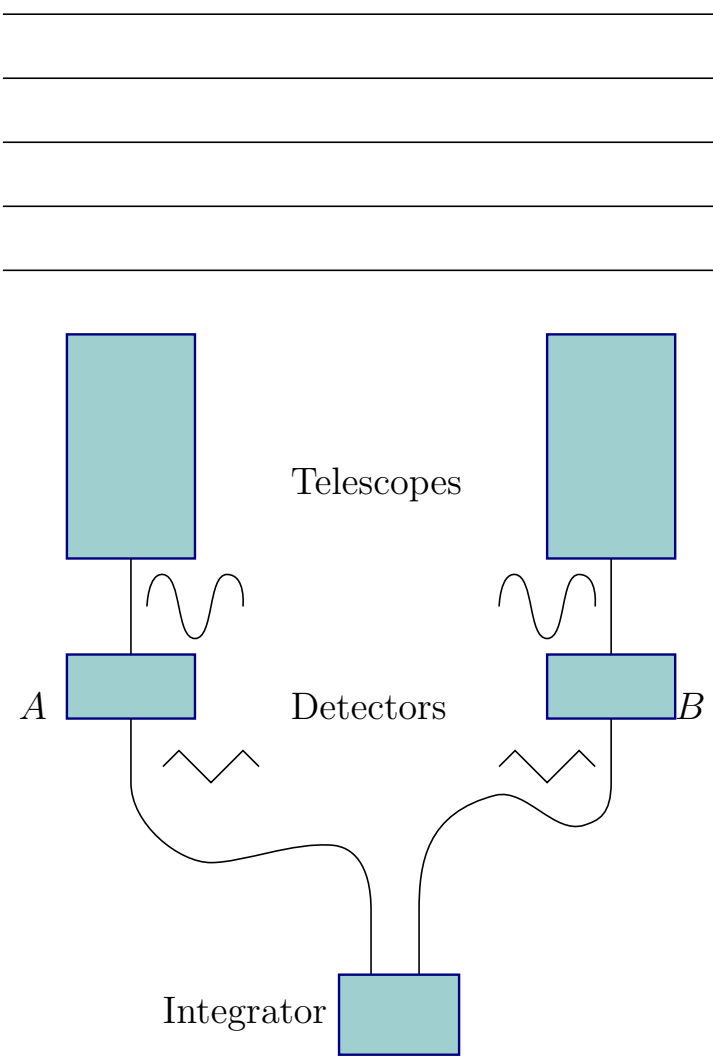

Figure 5.1: Light from a distant source is collected by two telescopes and directed at two detectors which each send a filtered signal to the integrator. Note the optical transmission of the signal is denoted by a curvy wave and electronic transmission via a zig-zag wave.

ometer is measuring angular sizes in a very similar way to the amplitude interferometer described in section 2. Actually we are measuring the square of the visibility function, γ^2, which has the unfortunate result of losing the phase information. This means we cannot unambiguously reconstruct any image asymmetries directly from two apertures in a Intensity Interferometer.

A major problem that we encounter with attempting to measure this second order effect is that it is a fluctuation in a relatively large current. Extraction of the signal from the noise is the main drawback

$$\left(\frac{S}{N}\right) = c(d)T_0/N(T_0) = A\alpha n|\gamma_d(0)|^2\left(\frac{\Delta f T_0}{2}\right)^{1/2} \tag{5.6}$$

The plotting of the measured electronic correlation against separation of the detectors then allows us to investigate the correlation function as we would have with the strength of the modulation in section 2.2.

5.2 An updated intensity interferometer

To see if we can make use of the benefits of the Intensity Interferometer setup while lessening the impact of the negatives associated with the scheme we can look at the effect of introducing modern electronics to the system such as looked at by Camarata and Horch(Camarata and Horch 2011). Focus on those areas we can readily improve on the 1970's incarnation within Equation (5.6) tells us that increasing the electronic resolving power of the detector Δf increases the signal to noise ratio. Specifically

$$\left(\frac{S}{N}\right) \propto (\Delta f)^{1/2}. \tag{5.7}$$

At the Narrabri Observatory they achieved a Δf of approximately 100MHz, with modern electronics we can readily increase that by a factor of ten with available laboratory equipment. As such we would see an increase by a factor of three on the results previously generated. For example in practical terms this means we could assess the radii of stars to a known percentage in a third of the time previously, or that we could assess three times as many objects in a given time. Further we can see that the signal to noise is also directly proportional to the systems quantum efficiency, that is how well we can turn incident photons into output electrons.

$$\left(\frac{S}{N}\right) \propto \alpha. \tag{5.8}$$

With a modern chain of electronics and detectors we can be optimistic of an efficiency of %50(Le Bohec and Holder 2006), an improvement of two and half times on the benchmark that was Narrabri.

Taken together these improvements give a signal to noise increase of seven over the Intensity Interferometer operated by Hanbury Brown. There they were limited to a magnitude of 2.5, with this improvement we can calculate the magnitude reachable by an improved, modern interpretation. Starting with equation (5.6) and rearranging for n, and

Characteristic	*Value*
Aperture	$4 \times 10^{-6}\text{m}^2$
Quantum efficiency α	0.70
Detector bandwidth	1GHz
Integration time	3600s

Table 5.1: Values for a theoretical Intensity Interferometer at the University of London observatory (ULO).

substituting the known relation between the photon flux and the magnitude

$$n = 2.512^{m_\lambda} F_0 \tag{5.9}$$

where m is magnitude and F_0 the photon flux of a zero magnitude star, $5\times10^{-5}\text{s}^{-1}\text{m}^{-2}\text{Hz}^{-1}$, we find

$$m_\lambda = -2.5 \log \left[\frac{(\text{S/N})}{F_0 A \alpha |\gamma(d)|} \sqrt{\frac{2}{\Delta f T_0}} \right]. \tag{5.10}$$

Aiming for a signal to noise of 5 and using apertures of 30m^2, an integration time of $T_0 = 3600\text{s}$ and values of α and Δf as suggested above, we can probe stars down to a magnitude of 6. A very significant improvement on previous instruments.

Alternatively we might consider a modest field setup at a small observatory such as the University of London Observatory at Mill Hill in North London. Here they are a pair of matched C14 Celestron Schmidt Cassegrain telescopes at a fixed baseline of approximately 7m. The important values for calculating the signal to noise are contained in table 5.1, with values for the detectors found in section 5.3.1.

Using these values we can plot the deepest magnitude reachable for a given integration time. In figure 5.2 each curve represents the variation in signal to noise ratio given an increase in integration time for varying magnitudes. At the bottom of the figure we can see the curve for a 0 magnitude star and each line gives the ratio with the magnitude increasing by one, up to $m = 6$, towards the top of the figure.

Taking reasonable values for the signal to noise and integration time, 5 and 10^3s respectively we can expect to be able to measure stars down to third magnitude even with the modest setup at ULO.

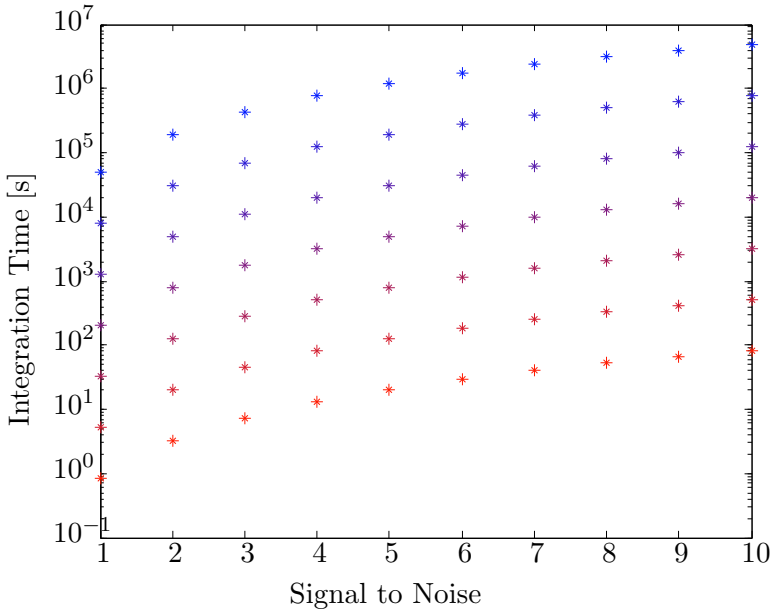

Figure 5.2: Integration time (s) vs Signal to Noise.

5.3 The lab model

To validate the improvements suggested in section 5.2 a simple lab top demonstrator was designed and implemented within the Optical Science Laboratory of the Astrophysics group at University College London. A major inspiration for this work was the second in a series of papers by Hanbury Brown and Twiss(Brown and Twiss 1958) that set out their initial proof of concept experiments carried out in the 1950's. The physical setup is shown in figure 5.3. This design recreates an interferometric setup on the labtop by placing a pinhole some distance from two detectors that the pinhole is not well resolved by either aperture[3]. To show that the coherence of the light in the aperture plane is being measured the detectors are firstly placed co-incident, and a non zero correlation is found and then separated by a distance that should show a zero correlation if the above theory is correct. To allow for the detectors to be placed co-incident as seen from the source a beam splitter is placed in inbetween the source plane and aperture plane. This is optically flat and has no effect other than redirecting 50% of the light at a right angle to the original optical axis. Thus optically the two detectors are at the same angle and distance from the source which implies $d = 0$ and we measure γ_0.

As seen in The Intensity Interferometer(Hanbury Brown 1974) we can calculate the

[3]See (5.12) for details.

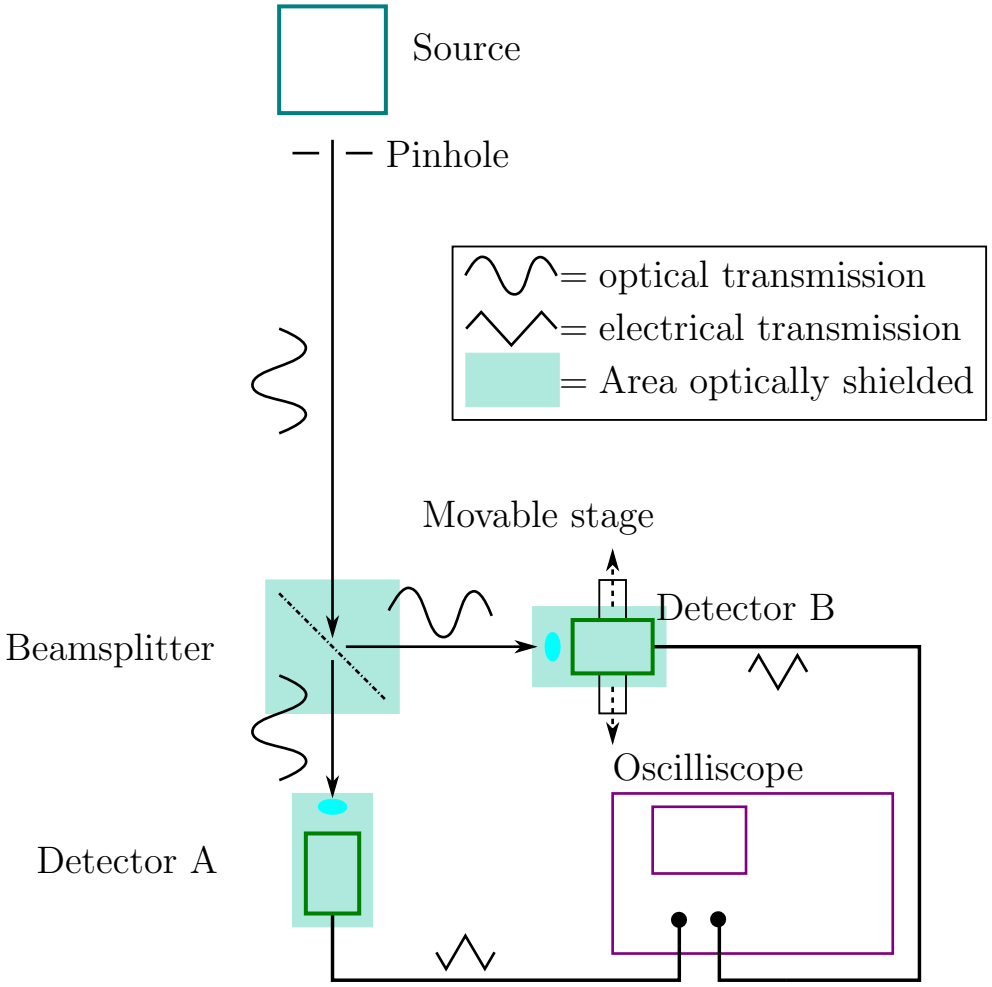

Figure 5.3: Diagram of the labtop Intensity Interferometer setup showing the optical setup. Light from the source via a pinhole hits the beam splitter and sent to two detectors such that they are optically coincident. A lens in front of each detector increasses the collection area. Detector B can move perpendicular to the optical axis simulating increasing the baseline separation of the detectors as in a functioning interferometer.

zero separation correlated signal $S(0)$ from

$$S(0) = ATb_\nu f\left(\frac{a_1\theta_1\pi}{\lambda_0}\right) f\left(\frac{a_2\theta_2\pi}{\lambda_0}\right) \int \alpha^2(\nu) n_0^2(\nu) \mathrm{d}\nu \tag{5.11}$$

This partial coherence factor, $f\,(\eta)$, is dependent on the product of the angular size of the source and linear size of the aperture, with the parameters defined in figure **??**. Where

$\eta \ll 1$ then $f(\eta) \to 1$, or

$$\begin{aligned} \text{for } \frac{a\theta}{\lambda} &= \eta \ll 1 \\ f(\eta) &\approx 1 \\ \text{for } \frac{a\theta}{\lambda} &= \eta \approx 1 \\ f(\eta) &\approx 0.7 \end{aligned} \tag{5.12}$$

From here a movement perpendicular to the optical axis has the effect of increasing d and at $d = D$ we would find a zero correlation. If we were in the far field then to calculate D we refer can back to equation (1.1) on page 15 and considering $L = 3$m, a pinhole at the source of 0.1mm^2, lenses at the detector stopped to 5mm and $\lambda = 500$nm so that $\theta = 3.33 \times 10^{-5}$radians and

$$\begin{aligned} D &\approx \frac{1.22 \times 500\text{nm}}{3.33 \times 10^{-5}} \\ D &\approx 1.8\text{cm} \end{aligned} \tag{5.13}$$

However because we are not in the far field, as seen $\eta \approx 1$, we need to consider the effects of only partial coherence over the aperture and as such the correlation S is actually measuring the overlap of the two apertures. So a movement to a separation of $D = 2 \times$ aperturewidth should result in a zero correlation, ie at

$$d = D \qquad\qquad S(d) = 0 \tag{5.14}$$

Which for our apertures of 5mm width suggests a separation of 1cm.

5.3.1 The detectors

The detectors used are a pair of avalanche photo diodes (APD) from Hamamatsu Photonics, Japan. These were provided in a module with a built in amplifier, unit C5658. The unit is biased to improve the sensitivity and within this circuit is a temperature control to moderate the bias voltage to appropriate levels. The technical details are summarised in table 5.2. The detector is well matched in terms of response time to the signal frequency of a thermal source. As seen in chapter 1 for a thermal source this is expected to be a

Characteristic	*Value*
Bandwidth	50kHz to 1GHz
Wavelength	400nm to 1μm
Sensitivity ($\lambda = 800$nm)	2.5×10^5 V/W
Maximum input	10 mW
Amplifier Gain, B	100

Table 5.2: Important Characteristics as per APD module C5658 instruction manual.

frequency around 1GHz and with a wavelength response wide enough to detect a source out to near infrared frequencies.

The output of the detector can be calculated from

$$V_{\text{out}} = P_{\text{in}} \times \text{radiant system} \times \text{diode gain} \times \text{load resistance} \times \text{amplifier gain}. \quad (5.15)$$

For high speed optical signals the lower bound to the sensitivity is determined by the dark noise present in the semi conductor. For a low noise amplifier the signal to noise is improved by an amount approximately equal to the amplifier gain B.

5.3.2 The electronics

For the integrator an oscilloscope was used to collect the output from the detectors and to save the data to file. This file is later analysed in software to determine the correlation of the detector outputs.

The oscilloscope used was a GW Instek GDS-2302A with a Bandwidth of 300MHz which under samples the detectors at 1GHz as seen in table 5.2. This is the limiting bandwidth of the system. The sensitivity at 100MHz is approximately 100mV which via equation (5.15) determines a minimum light intensity of

$$\begin{aligned} P_{\text{in}} =& V_{\text{out}} \times \frac{1}{\text{radiant sensitivity} \times \text{diode gain} \times \text{load resistance} \times \text{amplifier gain}} \\ P_{\text{in}} =& \frac{100\text{mV}}{0.5\text{A/W} \times 100 \times 50\Omega \times 100} \\ P_{\text{in}} =& 2.5 \times 10^{-6} = 2.5\mu\text{W} \end{aligned} \quad (5.16)$$

Which dictates a bright source.

To record the data from the oscilloscope use was made of the *Datalog* function. This is not native to the equipment but was downloaded and installed as an additional function. This function allowed the waveform as currently displayed to be saved to memory and

then downloaded to a memory stick. This could be repeated for a set time, the sum of which would be the total integration time. This leaves a set of files to be analysed as described in section 5.3.4.

5.3.3 The source

The theory of Intensity Interferometry requires a thermal source that comprises a very large number of independent emitters so that the source is a collection of varying intensity. For scientific objects of interest it is not unrealistic that we can expect such a thermal source. For the laboratory experiment an incandescent bulb was used. This was imaged onto a pinhole to provide a thermal source with known size and shape.

We can calculate the photon flux at the detector with some fair assumptions of the physical parameters present in the system.

Beginning with the source, an incandescent lamp, we can measure the input power via

$$\begin{aligned} P &= IV \\ P &= 650\text{mA} \times 12\text{V} \\ P &= 7.8\text{W} \end{aligned} \tag{5.17}$$

Where P is the electrical power through the lamp, I and V the current through and voltage across the filament. And so we expect our output to be in the region of 10 watts.

To calculate the output power of the lamp we turn to the blackbody equation

$$I = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{kT\lambda}\right) - 1} \tag{5.18}$$

Where h is Planck's constant, c is the speed of light, λ the wavelength at which the power is emitted, k is the Stefan Boltzmann constant and T the temperature of the filament.

Integrating this over the region of the spectrum that the detector is sensitive to will gives us a relevant power output. Allowing for losses due to glass surfaces, the light dispersing over a hemisphere centered on the pinhole in front of the lamp, the emmisivity of the tungsten at 15% and the detector efficiency we can see that the total input power comes out at 3.0×10^{-6} watts. Assuming an average wavelength of 500nm we have a photon flux of 7×10^{24} photons per nano second of integration time. Which matches the requirements of equation (5.16)

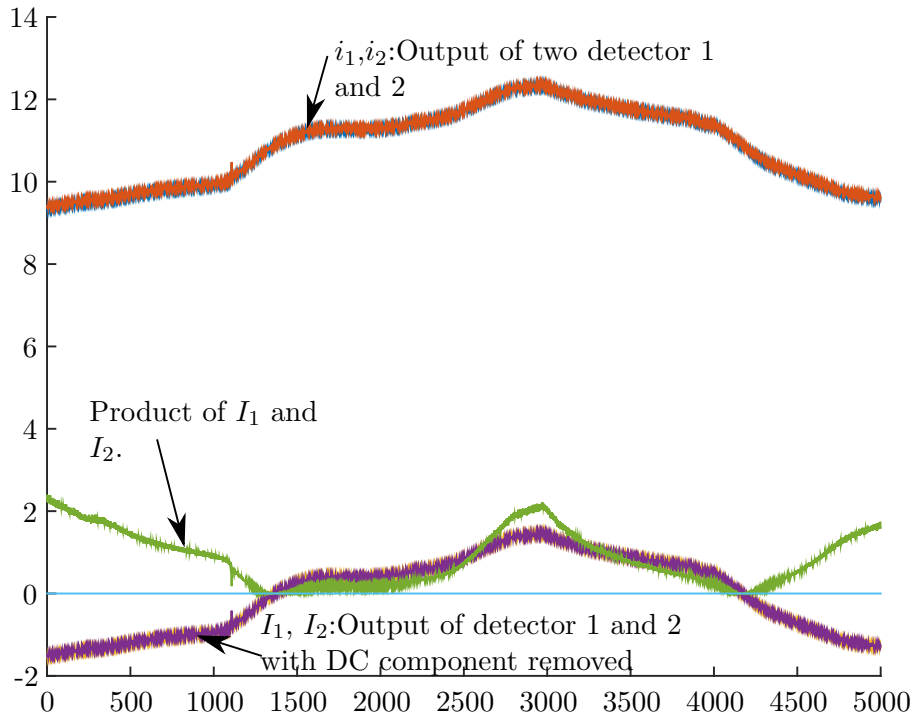

Figure 5.4: A simulated recording of an electronic signal generated by a detector. The top trace shows the raw output of a time varying, AC intensity about some DC constant. The bottom trace shows how the AC component after removal of the DC offset. The central trace shows the product of the two AC traces over time. Note for this highly correlated sample this correlation is "high".

5.3.4 The software

The data from each channel was simultaneously recorded by the oscilloscope and transferred to a pc running Matlab for analysis. This software calculates the average value of the light over the total integration time, then calculates the correlation value from equation (5.6).

Figure 5.4 shows the output of Matlab after inputting the data. It shows three traces using simulated data. The top traces show i_A and i_B which represent the electronic output from two detectors, A and B as seen in figure 5.3. These traces are the combination of the AC and DC components inherent the incident light. The bottom trace is the AC trace alone with the DC trace removed. These AC traces are then multiplied together timewise to produce the middle trace. The correlation as defined in (5.11) is equal to the sum of this trace for the correct integration time. For the near perfect correlation in figure 5.4 sums to around an, arbitrary, value of around 4200. In figure 5.5 we have input a delay

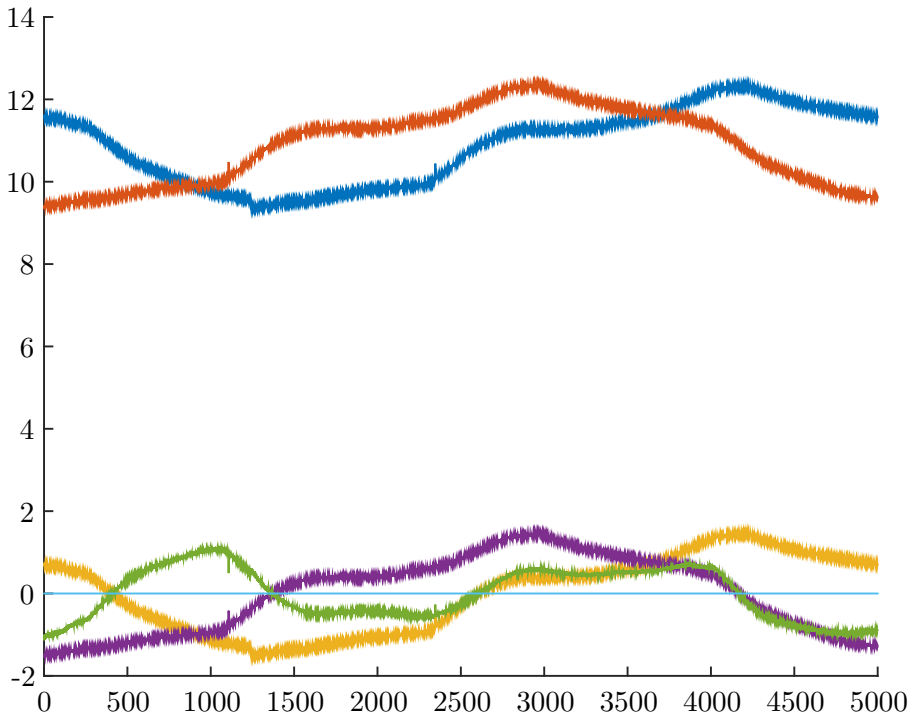

Figure 5.5: Here the simulated data as used in figure 5.4 has a time offset applied to represent a path length difference. The red trace is at zero path length and the blue trace has an offset of about 1200 points. The purple and yellow trace show the input signals with the DC component removed. Green shows the product of the AC signals, note how it now goes negative for parts of the integration time.

into one signal so that we see a much lower correlation in the original signals. Importantly we note that the product trace varies both positive and negative and that the sum of this therefore much closer to zero at around 15.

To obtain the DC component the arithmetic mean is taken of the trace and this is subtracted piecewise from each data point to leave the AC signal alone. This simulated data shows a much larger AC signal than would be expected in a thermal source.

5.3.5 Results

The results taken from the lab were effectively a null result. Six runs were made both at zero displacement and with the detectors offset to 8mm. These results are summarised in table 5.3

The low measured value of Γ fails to confirm the expected result of a positive value. This is possibly due to the short integration time. While the system was running for the

Table 5.3: Table of runs conducted on labtop Intensity Interferometer.

	Equipment Setup		
Run Number:	d	T	Γ
1	0	5	0.001
2	0	30	0.002
3	0	35	0.013
4	0	35	0.002
5	0	35	-0.0086
6	8mm	35	-0.0121

full time, there was a latency within the oscilliscope recording method whereby data was only being recorded for a short period of this time. As such this needs further investigation and improvement, else the low actual integration time will ultimately limit the signal to noise ratio.

5.4 A space based Intensity Interferometer

This section will try to describe a space mission based on Intensity Interferometry and consider what can be achievable with a low cost and high reliability mission. It is considered to be an alternative mission paradigm, distinct from those mentioned in chapter 1 that have repeatedly failed to see the light of day or indeed night.

5.4.1 Architectural Overview

An Intensity Interferometer can be constructed from just five parts, two apertures, two receivers and a correlator. In practice we need to add ancillary items to direct the light and electrical signals to where they are needed. These parts can be considered independently of the main parts however so here we will concentrate on considerations directly related to these main parts. As with any space mission the total mass of the system is the primary cost and as such the fewer parts the better. However the technical capabilities of the system are much improved on increase from the basic five elements. With an interferometer there is a tension between spreading your apertures far and wide to increase the sensitivity to longer wavelength fourier components[4] which correspond to smaller details on the source with the ability to substantially measure the intermediate frequencies. The measurement of these lower frequencies relies on being able to "fill" the aperture plane with receivers over the course of an observation. This may be done by rotation of the array of apertures

[4]For further details on fourier optics see section 4.1.

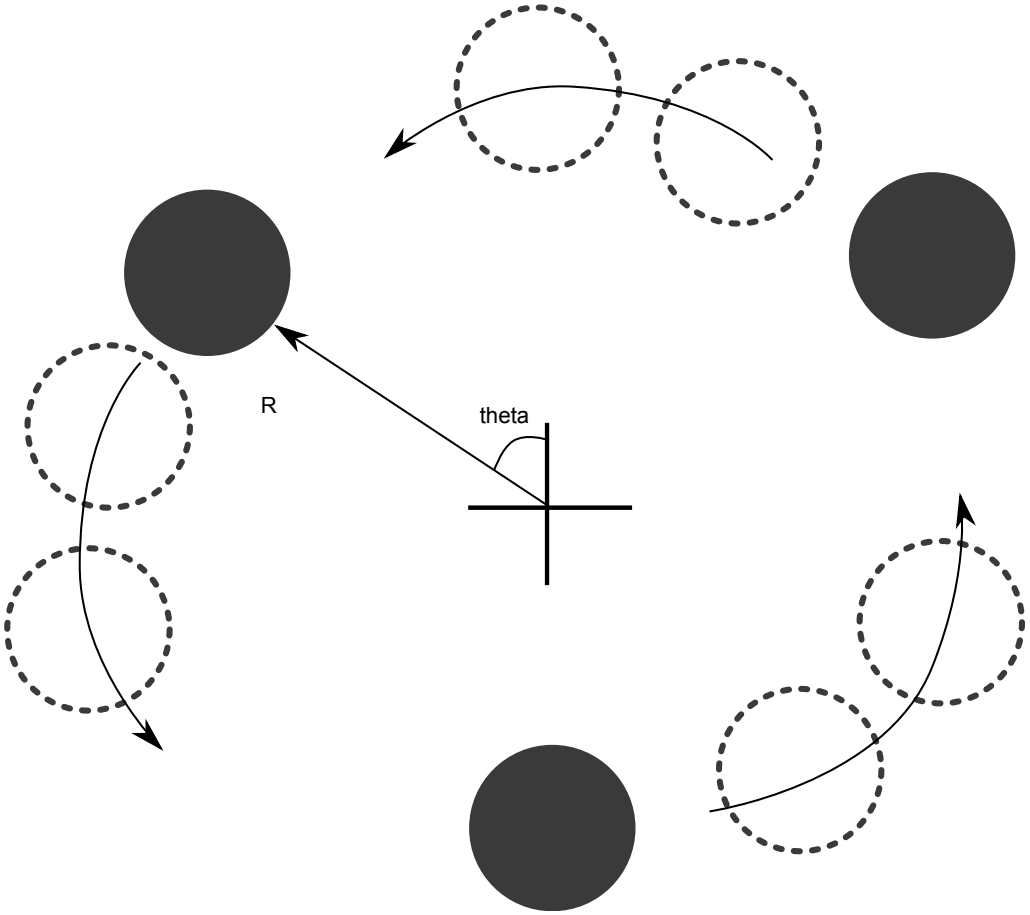

Figure 5.6: As the aperture array rotates about θ the sky is filled with footprints of the apertures. Varying R allows for full coverage. In practical terms it is not necessary to have 100% coverage and integrationtimes may limit the amount of time spent in any one configuration.

as shown in figure 5.6. These considerations are discussed in sections 5.4.3 and 5.4.4. With the low cost of this scheme the mission proposal would be for a core of five to ten apertures able to image objects of large angular size such as the planetary orbits in table3.4 and three or four "outriggers" flying at a large separation able to measure objects of low angular size such as stellar discs. This suggests a total number of spacecraft in the region of $5+3+1=9$ to $10+4+1=15$, each aperture is envisaged to be 1m in diameter.

The total error budget in the system is set by choosing an arbitrary limit for loss of coherence due to path length mismatches between the source and the apertures. A reduction in coherence by 50% allows a path mismatch of around 30cm[5] and therefore the optical surfaces can be of the "light bucket" type. This can significantly reduce the cost of manufacture and eases concerns over correct figuring of the mirror in space. Formation

[5] As described on pages 148 and 149, An Introduction to Optical Stellar Interferometry(Labeyrie et al. 2006) and chapter 5.

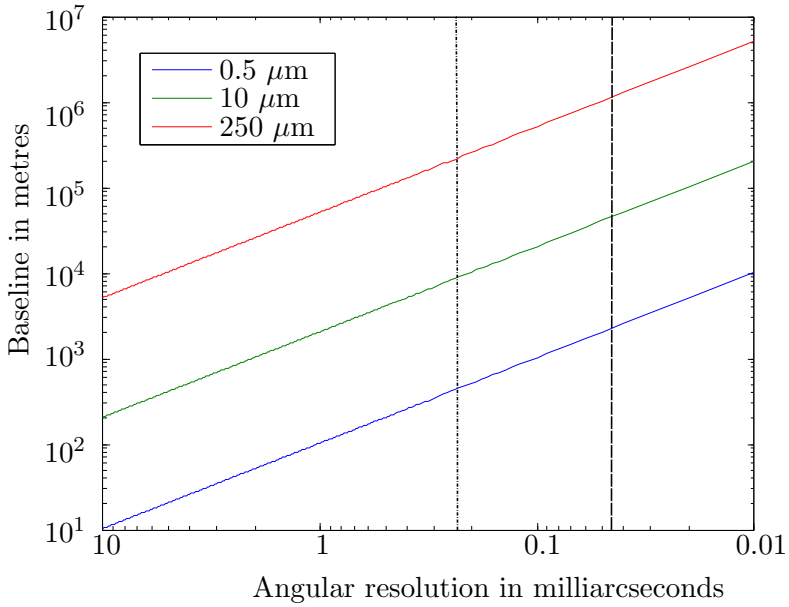

Figure 5.7: Baseline versus angular resolution for three wavelengths. The vertical dashed line represents the angular size of a G class star at 100pc and the dot-dash line the same star at 40pc.

flying concerns can also be lessened, maintaining position in formation flight is still a largely unproven technology. Lim(Lim and Ahn 2014) writes that position keeping to a millimetre accuracy is possible. Kristiansen and Nicklasson(Kristiansen and Nicklasson 2009) suggest a sub metre accuracy available after $\approx$ 5 minutes. This level of control is well capable of operating within the total error budget of the satellite array.

Each spacecraft would record the measured intensity for some integration time and then send this data to the central hub spacecraft for primary correlation.

5.4.2 Baseline versus angular resolution and field of view

As explained in chapter 2 the baseline of an interferometer may be seen as analogous to the diameter of a monolithic aperture when considering the finest detail that may be resolved. As such it is tempting to fly our apertures as far apart as reasonably possible in the aim of maximising our angular resolution. However the further apart the spacecraft fly the higher the demands on the station keeping of the apertures. Also an interferometer is only able to measure angular frequencies equal to $1/B$ and the overall picture is modelled from a variety of visibility measurements as seen in figure 2.3. Therefore some sophistication is required in the spacing and positioning of our apertures. Figure 5.7 shows angular resolution versus baseline annotated with the size of several objects of interest at three different wavelengths. We can see that at a distance of 100pc a baseline of at least 10^5m

is required to view in the FIR while only 2×10^2 in the visible. A problem of inverse wavelength dependency mentioned in chapter 1. Therefore for this mission proposal a baseline of 5×10^4 metres is suggested as this includes the ability to resolve G class stellar discs at a distance of 100 parsecs. The core would be 200 metres in size as this allows sizes as small as 0.5AU at 100pc to be resolved. This ability allows for imaging of of the habitable zone of 0.5 to 2AU around G class stars.

5.4.3 Apertures and phase closure

The advantage of a monolithic lens over the diluted nature of an interferometric pupil is its ability to simultaneously measure all fourier frequencies up to λ/D for some diameter, or baseline, D. With this information an image can be obtained instantly, for interferometric imaging some finite time is necessary while as many fourier components are compiled as wanted. We do not need to fully measure the pupil plane for a usable image, some partial measurement can be made and via modelling a usable image can be obtained. While the contrast of the visibility function can be readily obtained it's phase is much harder in a practical instrument. We can regain the phase information from the "closure phase"(Cornwell 1989)(Baldwin et al. 1986). It was first described by Jennison in 1958(Jennison 1958).

Closure phase works by comparing the phase from three apertures such that the original, "absolute" value is replaced with a relative value and the ambiguity within the image is lowered. Without an assumption about the original source intensity or with a completely filled aperture then ambiguity will never be zero but we can make substantial progress with only a partially filled aperture.

As described in Kraus et al., Infrared imaging of Capella with the IOTA closure phase interferometer(Kraus et al. 2005) the closure phase can be used for the imaging of Binary stars, with the ability to plot angular separation as a function of time. Phase closure is necessary for high angular resolution of stellar discs, giving information about variations in surface temperature and equatorial bulging. Highly detailed images including dynamical processes can be reconstructed via the method such the Equatorial disc conjunction of ϵ-Aurigae(Kloppenborg et al. 2010), where the equatorial dust disc of one of the binary system obscures it's partner.

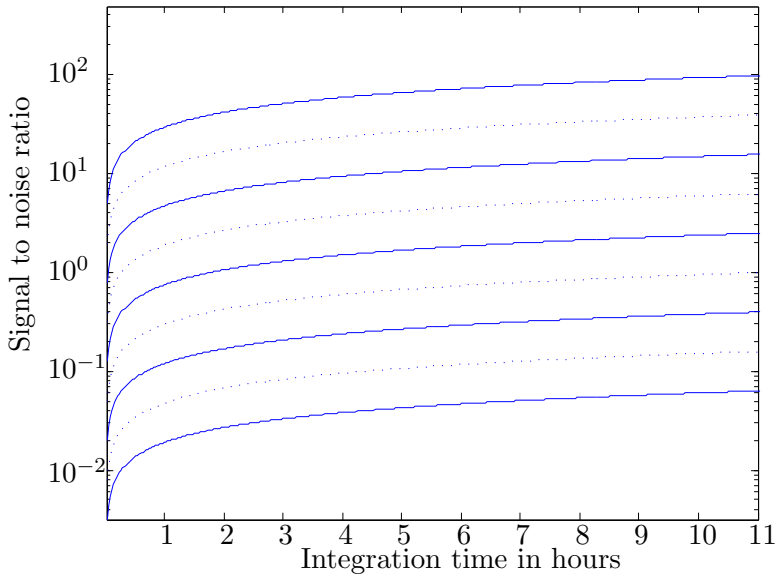

Figure 5.8: Signal to noise ratio for decreasing magnitude vertically. The lowest dashed line represents 8^{th}mag and the top line 0^{th}mag, each line between being a change in magnitude of one.

5.4.4 Observation time versus Signal to Noise

Section 5.1 references the low sensitivity of an Intensity Interferometer and therefore we are sensitive to maximising the integration time spent observing objects of interest. As can be seen in figure 5.2 an integration time of one hour allows a signal to noise ratio (snr) of five to be measured for a magnitude two star using modest equipment. With modern equipment and larger mirrors we can investigate the deepest magnitude we can reach with an Intensity Interferometer . With six metre sized mirrors, an efficiency of 0.5 and electronics operating in the region of 100MHz we can expect observation times of around one hour to reach a snr of around 1 for a 7^{th}mag star as shown in figure 5.8. Within 10 hours a magnitude 4 star can be measured to a SnR of approximately 5. Being space based has the major advantage of allowing integration times of over 12 hours, for given directions on the sky. And as such much longer integrations can take place. Figure 5.9 shows the same function but plotted against a log time to see the effects of longer integration times. Due to the nature of the function there are ever diminishing returns on integration times. Setting an arbitrary limit at 100 hours we can hope to achieve a signal to noise ratio of unity down to a magnitude of nine and ten can achieved in a little over 100 hours. Using the parameters in table 5.4 this sets our limiting magnitude at a conservative nine. Using this magnitude of nine we can determine a distance at which we can hope to measure a G class star. Taking our own star as a prototype the absolute

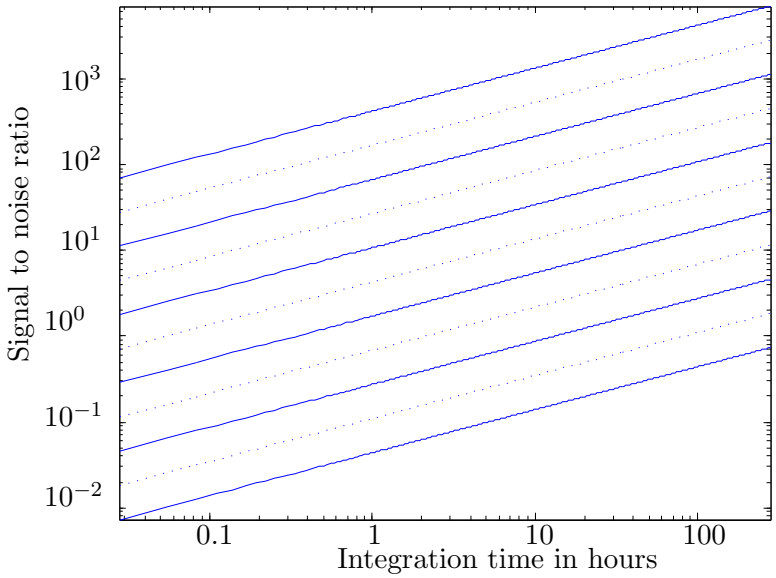

Figure 5.9: Plotting time on a log scale shows that for very long exposures of 100 hours we can reach a signal to noise of one for at most a magnitude nine star. This could be pushed to magnitude ten for a little extra time. The lowest solid line represents a 10^{th}mag star and the top line 0^{th}mag star, each line between being a change in magnitude of one.

Parameter	Value
Collecting area, A	30m^2
Overall system efficiency, η	0.5
Electronic bandwidth, ΔT	100 MHz
Integration time, T	

Table 5.4: Parameters of a space mission.

magnitude of such a star is five and using the distance modulus relationship,

$$\begin{aligned} m - M &= 5 \log d - 5, \\ d &= 10^{\frac{M-m+5}{5}}, \end{aligned} \tag{5.19}$$

we find that,

$$\begin{aligned} d &= 10^{\frac{10-5+5}{5}} \\ d &= 10^2 \text{pc}. \end{aligned} \tag{5.20}$$

Here, m is the apparent magnitude, M the absolute magnitude and d, the distance in parsecs.

So with the parameters as in table 5.4 we can hope to reach close to a distance of 100 parsecs in line with our scientific goals.

This page was intentionally left blank

Chapter 6

Conclusions

Considering the technical capabilities of the instruments studied we can reach some tentative conclusions that an interferometric array may be some minor benefit to the study of exoplanets.

Of the three iterations of the instrument we have found that pupil masking is unlikely to be of sufficient benefit to overcome the problems of flux ratio and diffraction wing suppression. A large scale space based interferometer may well be able to probe some of the search space needed to advance the science of the field.

The flux ratio improvement calculated in chapter 4 is of the order of 10^2, well short of the required 10^7 as defined in chapter 3. As such the designs as presented would not be of sufficient efficiency to allow separation of the light of the host star from the exoplanet in orbit. Improvements in the fidelity of the coding could possibly enhance our analysis of the masks calculated as the resolution of the software is such that a lot of noise from the discrete fourier transforms. This increase of fidelity would require an increase in the processing power available.

The angular resolution of a space based Intensity Interferometer of sub 0.1 milliarcsecond into the far infrared with a baseline around 10^4 would allow significant numbers of stars to be studied with unprecedented accuracy. Integration times of a few hours are highly acceptable within a space mission where the Earth's rotation and diurnal nature are not relevant.

Ultimately this system would be a cheap and reliable first space based interferometer

from which other more adept machines can learn from and build upon. As seen in section 5.3 a working laboratory demonstrator could be made with simple cheap components. Though a refined model of this is required to show it's potential. A larger scale model could then be made to further understand the issues of such an instrument operating in a space environment.

Appendix A

The Van Cittert-Zernike Theorem

The Van Cittert-Zernike Theorem (VCZ Theorem) is the physical description of light propagation and the resultant coherence of two points in space-time[1]. As such it can used to construct a notion of a coherence area, the size of which can give information about the source responsible for emitting the light.

To begin with we define a 4D co-ordinate system of (X, Y, Z, t), that is three spatial and one time component. We take Z as parallel to a line connecting the centres of our three planes of interest, Σ, $\mathcal{A}$ and $\mathcal{F}$ and thus can be taken as the optical axis of our system. We will consider the general case where the system is 'stationary' and as such the behaviour of the electric field at a specific time E_t is time independent. That is while it will have an instantaneous random value, the character of the fluctuations remain the same over time and therefore a time average of appropriate length ($t >> 1/\nu$, where ν is some typical frequency) will give the same answer regardless of any offset in t. Assuming a plane parallel geometry of the source, σ, our apertures, $\mathcal{A}_1$, $\mathcal{A}_2$ and the focal plane $\mathcal{F}$ allows us to reduce the co-ordinate system to 2D, (X, Y). To avoid confusion (X, Y) will be mapped to three different notations, (l, m) in Σ, (u, v) in $\mathcal{A}$ and (x, y) in $\mathcal{F}$ as shown in figure 2.1.

The VCZ theorem can be summarised as follows, the intensity distribution in $\mathcal{A}$ is proportional to the fourier transform of the intensity distribution in Σ given that the

[1]Note that use of the word space-time is not to invoke notions of relativity but simply to refer to a particular point in space at a particular point in time.

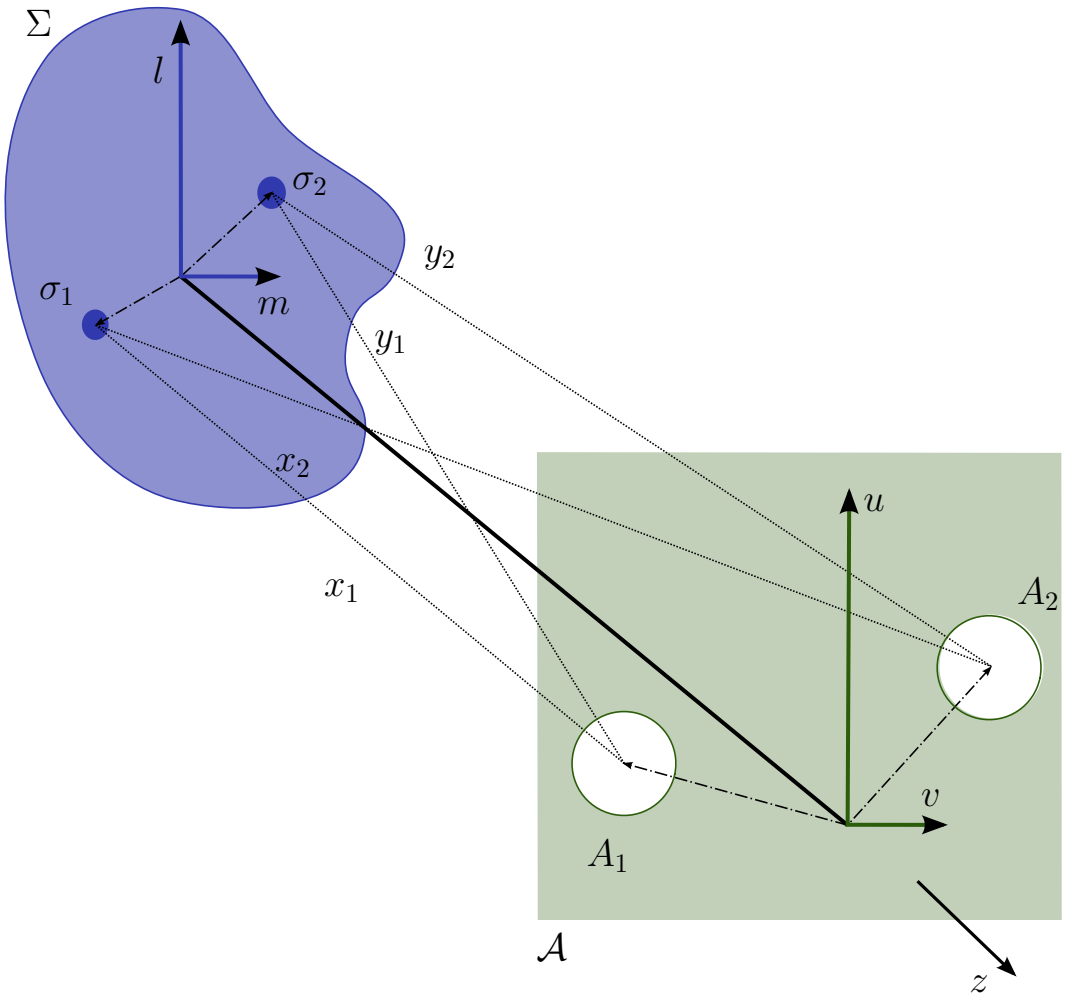

Figure A.1: We consider beams of light travelling from Σ to $\mathcal{A}$, from the point on the surface of Σ, σ to 2 apertures in $\mathcal{A}$.

distance between Σ and $\mathcal{A}$ is great enough that we might consider to be in the fraunhofer regime. To see this we begin by defining the situation between Σ and $\mathcal{A}$ in figure A.1. Within this picture an elementary point is considered on the surface of Σ, σ and the direct path that point makes towards two apertures $\mathcal{A}_1$ and $\mathcal{A}_2$; Which gives two vectors $\boldsymbol{r}_{i\mathcal{A}_1}$ and $\boldsymbol{r}_{i\mathcal{A}_2}$.

From this situation we can describe the disturbances at $\mathcal{A}_1$ and $\mathcal{A}_2$ as

$$V_{\mathcal{A}_1}(t) = \sum_i V_{i1}(t) \text{ and} \tag{A.1}$$

$$V_{\mathcal{A}_2}(t) = \sum_i V_{i2}(t) \tag{A.2}$$

with V_i the disturbance at σ_i.

Defining the *mutual intensity* $J(\mathcal{A}_1, \mathcal{A}_2)$ as

$$\begin{aligned} J(\mathcal{A}_1, \mathcal{A}_2) =& \langle V_{\mathcal{A}_1}(t) V^*_{\mathcal{A}_2}(t) \rangle \\ =& \sum_i \langle V_{i\mathcal{A}_1}(t) V^*_{i\mathcal{A}_2}(t) \rangle + \sum_{i \neq j} \sum \langle V_{i\mathcal{A}_1}(t) V^*_{j\mathcal{A}_2}(t) \rangle, \end{aligned} \tag{A.3}$$

and assuming a physical source of mutual incoherent oscillators then

$$\sum_{i \neq j} \sum \langle V_{i\mathcal{A}_1}(t) V^*_{j\mathcal{A}_2}(t) \rangle = 0 \tag{A.4}$$

for $i \neq j$. Now with reference to the vectors in figure A.1, $\boldsymbol{r}_{i\mathcal{A}_1}$ and $\boldsymbol{r}_{i\mathcal{A}_2}$, we can write the disturbances at $\mathcal{A}_1$ and $\mathcal{A}_2$ due to σ_i are respectively

$$\begin{aligned} V_{i\mathcal{A}_1}(t) =& A_i \left(t - \frac{\boldsymbol{r}_{i\mathcal{A}_1}}{v} \right) \frac{e^{-2\pi i \bar{\nu}(t - \boldsymbol{r}_{i\mathcal{A}_1}/v)}}{\boldsymbol{r}_{i\mathcal{A}_1}} \text{ and} \\ V_{i\mathcal{A}_2}(t) =& A_i \left(t - \frac{\boldsymbol{r}_{i\mathcal{A}_2}}{v} \right) \frac{e^{-2\pi i \bar{\nu}(t - \boldsymbol{r}_{i\mathcal{A}_2}/v)}}{\boldsymbol{r}_{i\mathcal{A}_2}}, \end{aligned} \tag{A.5}$$

$|A_i|$ represents the is the magnitude of the radiation from σ_i and arg A_i the phase, v is the velocity of light in the intervening medium between Σ and $\mathcal{A}$. Combining equations (A.3) and (A.5) gives us the coherence at two points on $\mathcal{A}$ due to a single point on σ,

$$\begin{aligned} \left\langle V_{i\mathcal{A}_1}(t) V^*_{i\mathcal{A}_2}(t) \right\rangle &= \left\langle A_i \left(t - \frac{\boldsymbol{r}_{i\mathcal{A}_1}}{v} \right) A^*_i \left(t - \frac{\boldsymbol{r}_{i\mathcal{A}_2}}{v} \right) \right\rangle \frac{e^{-2\pi i \bar{\nu}(\boldsymbol{r}_{i\mathcal{A}_1} - \boldsymbol{r}_{i\mathcal{A}_2})/v}}{\boldsymbol{r}_{i\mathcal{A}_1} \boldsymbol{r}_{i\mathcal{A}_2}} \\ &= \left\langle A_i(t) A^*_i \left(t - \frac{\boldsymbol{r}_{i\mathcal{A}_2} - \boldsymbol{r}_{i\mathcal{A}_1}}{v} \right) \right\rangle \frac{e^{-2\pi i \bar{\nu}(\boldsymbol{r}_{i\mathcal{A}_1} - \boldsymbol{r}_{i\mathcal{A}_2})/v}}{\boldsymbol{r}_{i\mathcal{A}_1} \boldsymbol{r}_{i\mathcal{A}_2}}. \end{aligned} \tag{A.6}$$

In astronomical applications the path length $\boldsymbol{r}_{i\mathcal{A}_2} - \boldsymbol{r}_{i\mathcal{A}_1}$ can be considered small compared to the coherence length of the light then $(\boldsymbol{r}_{i\mathcal{A}_2} - \boldsymbol{r}_{i\mathcal{A}_1})/v$ is negligible and from equations(A.3) and (A.6) we arrive at

$$J(\mathcal{A}_1, \mathcal{A}_2) = \sum \langle A_i(t) A^*_i(t) \rangle_i \frac{e^{2\pi i \bar{\nu}(\boldsymbol{r}_{i\mathcal{A}_1} - \boldsymbol{r}_{i\mathcal{A}_2})/v}}{\boldsymbol{r}_{i\mathcal{A}_1} \boldsymbol{r}_{i\mathcal{A}_2}} \tag{A.7}$$

A real physical source will comprise so many individual elements σ_i that it can be considered continuous and the sum in equation (A.7) becomes an integral. This also permits us to rewrite $\langle A_i(t) A^*_i(t) \rangle$ as $I(\sigma)$, the intensity per unit area of Σ. We can also reconsider our vectors as typical of the path from some source point to our apertures and lose the i subscript. Further refinement can be made with the use of the wave number $\bar{k}$

in place of $2\pi\bar{\nu}/v$ and so (A.7) becomes

$$J(\mathcal{A}_1, \mathcal{A}_2) = \int I(\sigma)_\Sigma \frac{e^{\pi i \bar{k}(\boldsymbol{r}_1 - \boldsymbol{r}_1)}}{(\boldsymbol{r}_1 \boldsymbol{r}_1)} \mathrm{d}\sigma. \tag{A.8}$$

Equation (A.8) is more useful when normalised to the overall brightness of the source which allows easy reference between different sources. We do this by dividing by the brightness of the source as seen at $\mathcal{A}_1$ and $\mathcal{A}_2$ leading to

$$j(\mathcal{A}_1, \mathcal{A}_2) = \frac{1}{\sqrt{I(\mathcal{A}_1)}\sqrt{I(\mathcal{A}_2)}} \int_\Sigma I(\sigma) \frac{e^{\pi i \bar{k}(\boldsymbol{r}_1 - \boldsymbol{r}_1)}}{(\boldsymbol{r}_1 \boldsymbol{r}_1)} \mathrm{d}\sigma \tag{A.9}$$

which is the formal van Cittert Zernike theorem. It tells us that the correlation between two points illuminated by a source will vary as separation increases and this variance depends on the size of the source and its distance.

Rewriting the vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in terms of the surface coordinates on Σ, l, m and $\mathcal{A}$, u, v, (A.9) can be rewritten into a very familiar form

$$j_{\mathcal{A}_1 \mathcal{A}_2} = \frac{e^{i\phi} \iint_\Sigma I(l, m) e^{-i\bar{k}(pl + qm)} \mathrm{d}l \mathrm{d}m}{\iint_\Sigma I(l, m) \mathrm{d}l \mathrm{d}m}. \tag{A.10}$$

This clearly resembles a fourier transform, so that a plot of the coherence in $\mathcal{A}$ will be equal to the value in the fourier transform of the source distribution. Taking a simple circular source as a top hat function we find the transform to be a first order bessel function, fixing one aperture to the centre of this distribution and moving the second aperture radially outwards the correlation drops as a bessel function drops, reaching zero at distance of $\frac{0.61\bar{\lambda}}{\theta}$, where θ is the angular size of the source as seen from $\mathcal{A}$.

From here we can understand the principles of how an interferometer works, by varying the separation of the two collecting areas we are investigating the size and form of the coherence area and, at least in theory, after suitable sampling we can reconstruct the source distribution.

Appendix B

Petal mask code output examples

In this appendix a few examples of the output from the petal mask code are given to show the variation in the focal plane arising from parameter variation.

Table B.1 contains a summary of the parameters used for four mask designs and are representative of the full output from the code.

	Parameter			
Chart number:	r	h	N	ra
1	400	500	8	1500
2	400	650	8	1500
3	400	600	14	1500
4	400	550	10	1500

Table B.1: The table contains the parameters as defined in section 4.3.1 , page 54.

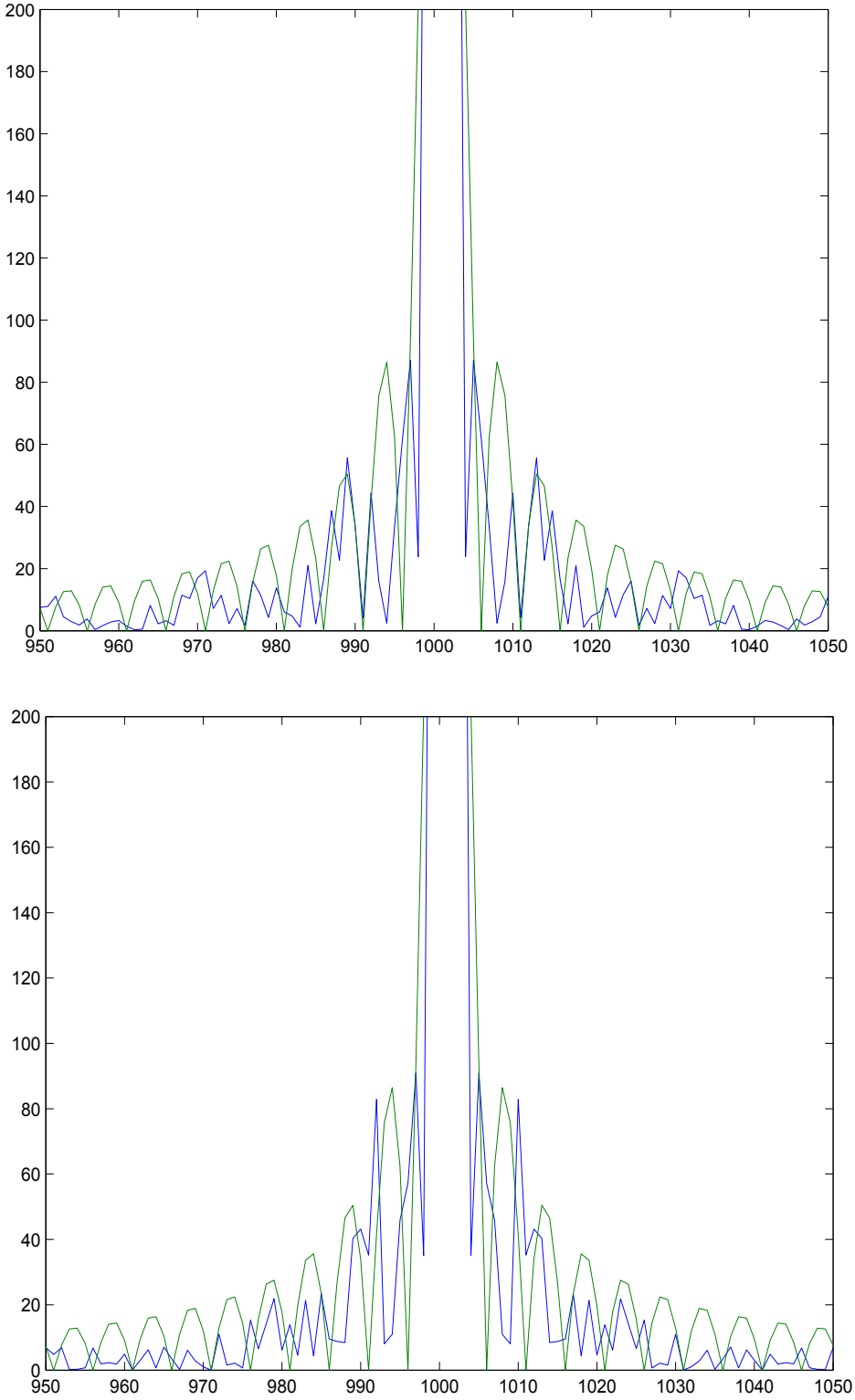

Figure B.1: Top: Chart 1. Bottom Chart 2.

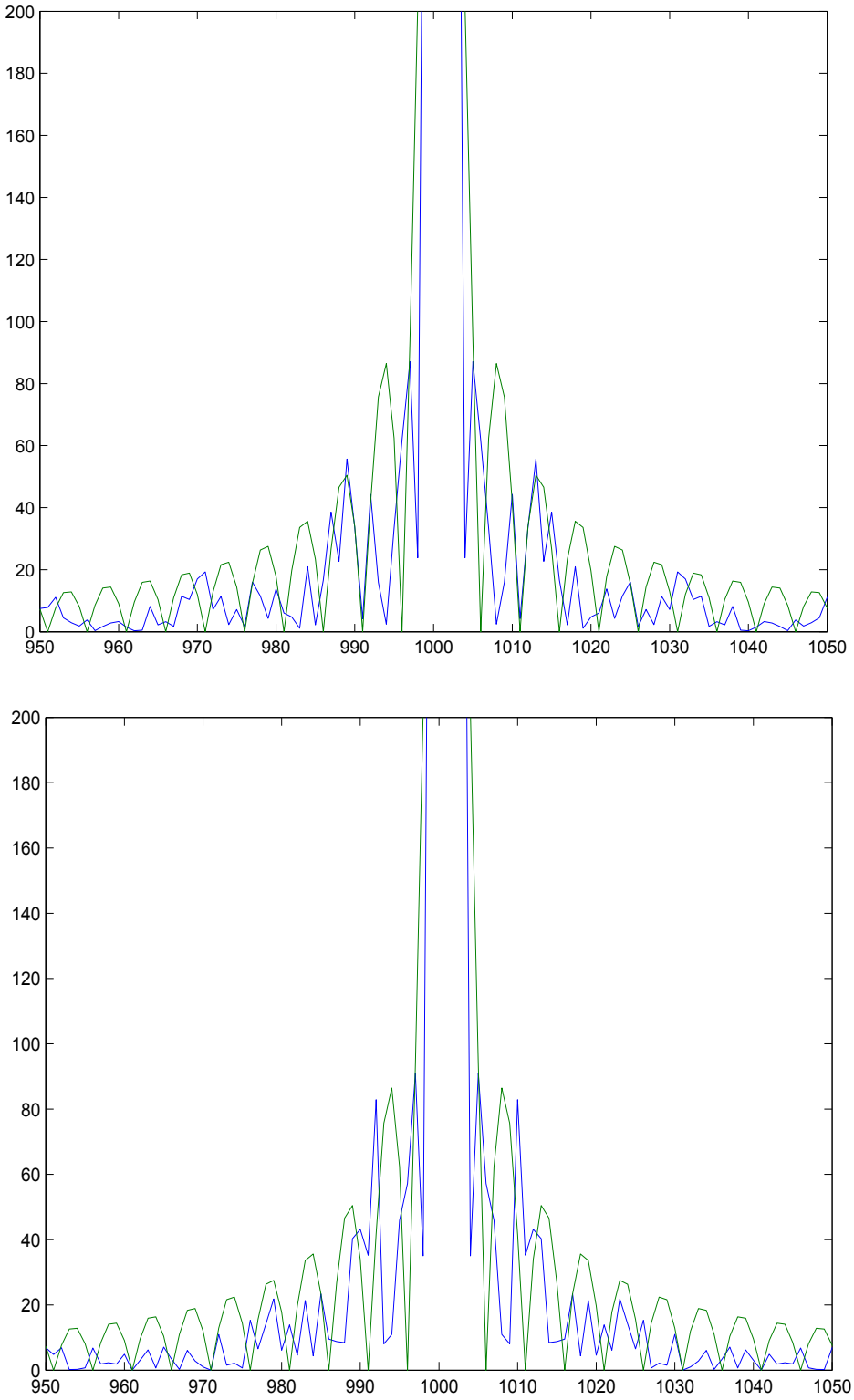

Figure B.2: Top: Chart 3. Bottom Chart 4.

This page was intentionally left blank

Appendix C

Starmask code example

```
clear all

%Mask Parameters
r = 400;                                    %inner radius
theta = linspace(0,2*pi(),500);             %angular position of spike vertex
theta2 = linspace(2*pi(),0,500);            %angular position of aperture
N = 10;                                      %number of spikes
M = 500;                                     %number of vertex points per spike
h = 500;                                     %height of spikes
%h1 = 300;                                    %mid-height of spikes
%h2 = 800;                                    %height of spikes from inner radius r
ra = 1500;                                   %radius of aperture
filler = NaN;
DetSize = 1500;
delta = 250;                            %fft resolution
edge = 15000;

%Guassian Constants   % f(x) = a.exp(-((x-b)^2)/(2.c^2))
Gsig = 1;
Gmu = 0;
```

```
Ga = 1/(Gsig*(sqrt(2*pi())));
Gb = Gmu;
Gx = 1;
Gc = Gsig;
%GaussFunc = a*exp(-((x-b)^2)/(2*c^2));

%vector pre allocation
x1 = zeros(5,N);
trianglex = zeros(5,N);
y1 = zeros(5,N);
triangley = zeros(5,N);

circley = r.*sin(theta + pi());
circlex = r.*cos(theta + pi());
circleay = ra.*sin(theta2+pi());
circleax = ra.*cos(theta2+pi());
%plot(x,y)

subY = linspace(-DetSize, DetSize, delta)'; %size of 'detector'/image
subX = linspace(-DetSize, DetSize, delta);

X = repmat(subX,delta,1);
Y = repmat(subY,1,delta);

for n = 0:N-1
    for m = 0:M/2;
    triangleVert1r(M*n+m+1) = r + (2.*m.*h/M);      %setting the spike points
                                                    %around the circle (r) (out)
    end
    for m = (M/2):M-1;
    triangleVert1r(M*n+m+1) = (r+h) - (2.*(m - M/2 ).*h/M);  %setting the spike
                                \%points (r) (back in again)
    end
```

```
end

for n = 0:N-1
    for m = 0:M/2
    triangleVert1theta(M*n+m+1) = [(pi()/2 + n.*(2*pi()/N))  -
        (Ga*6.249).*exp(-(((t riangleVert1r(M*n+m+1)./100) -
        (r/100)-Gb).^2)./(2.*Gc^2)).*(pi()/N*0.4)];   \%theta_0 + theta-n -
Theta_w + Theta_g
    end
    for m = (M/2)+1:M-1
    triangleVert1theta(M*n+m+1) = [(pi()/2 + n.*(2*pi()/N)) +
        (Ga*6.249).*exp(-((triangleVert1r(M*n+m+1)./100) -
        (r/100)-Gb).^2./(2.*Gc^2)).*(pi()/N*0.4)];  \%theta_0 + theta-n  +
Theta_g
    end
end
%calculating the vertex
%points in rectilinear coords
triangleVert1x = triangleVert1r.*cos(triangleVert1theta);
triangleVert1y = triangleVert1r.*sin(triangleVert1theta);

%x = cat(2, triangleA1x, triangleB1x, triangleC1x, triangleA1x);
%y = cat(2, triangleA1y, triangleB1y, triangleC1y, triangleA1y);
```

This page was intentionally left blank

Appendix D

Spergel Mask code example

```
clear all;
close all;

Factor1024 =2;
d = 1024*Factor1024;       %Resolution
R = 25;
xmin = -R;      %-R-(R/10);
xmax =  R;      %R+(R/10);
ymin = -30; %-R-(R/10);
ymax = 30;  %R+(R/10);


%perturbations of a, b and alpha
perta = rand/10;
pertb = rand/10;
pertalpha = rand/10;

a       = 0.4 * (1 + perta);  %aperture parameters
b       = 0.4 * (1 + pertb);
alpha   = 3.7 * (1 + pertalpha);
```

```
y_s = linspace(ymax,ymin,d)';   %vector 1 x d
Y_s = repmat(y_s,1,d);          % matrix d x d

x_a = linspace(xmin,xmax,d);                                % vector d x 1
y_a = a*R*(exp(-(x_a.*(alpha/R)).^2)-exp(-alpha^2));    %top edge; vector d x 1

x_b = linspace(xmin,xmax,d);                                % vector d x 1
y_b = -b*R*(exp(-(x_a.*(alpha/R)).^2)-exp(-alpha^2));   %bottom edge vector d x 1

%plot(x_a,y_a);
Y_a = repmat(y_a,d,1); % repeats topedge function for each row
Y_b = repmat(y_b,d,1); % repeats bottomedge function for each row
%the above produces two matrices, largely similar.  they repeat f(x) for
%each y row, as y(x) is not a function of y.

% imagesc(Y_s); colorbar
% imagesc(Y_a); colorbar
% imagesc(Y_b); colorbar

twodm = logical(Y_a > Y_s ) + logical(Y_b < Y_s)  - 1 ;
%binary 2D function f(x,y)
%This quries whether each pixel location is located within the open region
twoDM = twodm./100; %to reduce the influence of the aperture function.
figure(1);
imagesc(twoDM) ; axis off;

fftr = (fft2(fftshift(twoDM),1024*Factor1024,1024*Factor1024));
fftR = (fftr);    %Fourier transform of f(x,y)
fftZ = fftshift(fftR);
%manipulates output to put zero frequency in centre of plot
%normalZ = fftZ/max(fftZ);
```

```
figure(2);
imagesc(log(abs(fftZ))); axis off; colormap(hot); colorbar;
%figure(3);
%plot(fftZ(:,d/2))

par = [a, b, alpha];
%plot(x_a,y_a)
%hold on
%plot(x_a,y_b)

% flipx_b = flipud(x_b');
% flipy_b = flipud(y_b');
% allx = cat(1,x_a',flipx_b);
% ally = cat(1,y_a',flipy_b);
% output = cat(2,allx,ally);

%dlmwrite('spergelmask.uda',output,' ')

Fieldsize = 1024*2;
screen6464 = zeros(Fieldsize,Fieldsize);

Ysep = 30;
Xsep = 30;
starsize = 1;
starbrightness = 10^2;
starY1 =  Fieldsize/2+Ysep;     %Fieldsize/2 - starsize;
starY2 =  starY1 + starsize;    %Fieldsize/2 + starsize;
starX1 =  Fieldsize/2+Xsep;     %Fieldsize/2 - starsize;
starX2 =  starX1 + starsize;    %Fieldsize/2 + starsize;

size = 1;
brightness = 1;
```

```
PlanetY1 = Fieldsize/2-Ysep;
PlanetY2 = PlanetY1+size;
PlanetX1 = starX1; %512-Xsep;
PlanetX2 = starX2; %PlanetX1+size;

size2 = 1;
brightness2 = 1;
Planet2Y1 = starY1; %512;
Planet2Y2 = starY2;
Planet2X1 = Fieldsize/2-Xsep;
Planet2X2 = Planet2X1+size2;

% On axis Star
screen6464(starY1:starY2,starX1:starX2) = starbrightness;
% Off axis object 1
screen6464(PlanetY1:PlanetY2,PlanetX1:PlanetX2) = brightness;
% Off axis object 2
screen6464(Planet2Y1:Planet2Y2,Planet2X1:Planet2X2) = brightness2;



figure(4);
imagesc(screen6464);        %Star field

fftscreen = fft2(fftshift(screen6464),1024*Factor1024,1024*Factor1024);
fftScreen = fftshift(fftscreen);

figure(5);
imagesc(abs(fftScreen))     %FT of star field

Result = (twoDM).*(fftScreen)';  %Product of (Ap) and FT(Field)
figure(6);
imagesc(log(abs(Result)));
```

```
image = ifft2(fftshift(Result),1024*Factor1024,1024*Factor1024);
Image = fftshift(image);
figure(7);
imagesc(log(abs(Image))); colorbar;

closeupLOImage = (abs(Image(PlanetX1-50:PlanetX2+50,PlanetY1-50:PlanetY2+50)));
closeupLOMax = max(max(closeupLOImage));
figure(8)
surf(log(closeupLOImage));

closeupHIImage = (abs(Image(Planet2X1-50:Planet2X2+50,Planet2Y1-50:Planet2Y2+50)));
closeupHiMax = max(max(closeupHIImage));
figure(9)
surf(log(closeupHIImage));

closeupHI = (abs(Image(Planet2X1-10:Planet2X2+10,Planet2Y1-10:Planet2Y2+10)));
closeupLO = (abs(Image(PlanetX1-10:PlanetX2+10,PlanetY1-10:PlanetY2+10)));

averagehi = mean(closeupHI);
averageHi = mean(averagehi);
averagelo = mean(closeupLO);
averageLO = mean(averagelo);

ratioHI = closeupHiMax/averageHi;
ratioLO = closeupLOMax/averageLO;
```

This page was intentionally left blank

Bibliography

Alonso, R., Auvergne, M., Baglin, A., Ollivier, M., Moutou, C., Rouan, D., Deeg, H., Aigrain, S., Almenara, J., Barbieri, M. et al., 2008, *arXiv preprint arXiv:0803.3207*

Baldwin, J., Haniff, C., Mackay, C. and Warner, P., 1986

Born, M. and Wolf, E., 1999, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (CUP Archive)

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E. et al., 2010, *Science*, **327**(5968), 977

Bouchy, F., Mayor, M., Lovis, C., Udry, S., Benz, W., Bertaux, J.-L., Delfosse, X., Mordasini, C., Pepe, F., Queloz, D. et al., 2008, *arXiv preprint arXiv:0812.1608*

Brown, R. A. and Burrows, C. J., 1990, *Icarus*, **87**(2), 484

Brown, R. H. and Twiss, R., 1958, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, **243**(1234), 291

Brown, R. H. and Twiss, R. Q., 1957, *Royal Society of London Proceedings Series A*, **242**, 300

Brown, R. H. and Twiss, R. Q., 1958, *Royal Society of London Proceedings Series A*, **248**, 199

Camarata, M. A. and Horch, E. P., 2011, in *American Astronomical Society Meeting Abstracts 218*, volume 218 of *American Astronomical Society Meeting Abstracts*, pp. 409.04–+

Clampin, M., 2011, in *SPIE Optical Engineering+ Applications*, International Society for Optics and Photonics, pp. 814605–814605

Colavita, M., Akeson, R., Wizinowich, P., Shao, M., Acton, S., Beletic, J., Bell, J., Berlin, J., Boden, A., Booth, A. et al., 2003, *The Astrophysical Journal Letters*, **592**(2), L83

Colavita, M. M., Boden, A. F., Crawford, S. L., Meinel, A. B., Shao, M., Swanson, P. N., van Belle, G. T., Vasisht, G., Walker, J. M., Wallace, J. K. et al., 1998, in *Astronomical Telescopes & Instrumentation*, International Society for Optics and Photonics, pp. 776–784

Cornwell, T., 1989, *Science*, **245**(4915), 263

de Souza Jr, A. and de Souza Jr, A., 2002, *EAS Publications Series*, **5**, 95

Dravins, D., LeBohec, S., Jensen, H. and Nuñez, P. D., 2012, *New Astronomy Reviews*, **56**(5), 143

Fischer, D. A., Howard, A. W., Laughlin, G. P., Macintosh, B., Mahadevan, S., Sahlmann, J. and Yee, J. C., 2014

Foreman-Mackey, D., Hogg, D. W. and Morton, T. D., 2014, *arXiv preprint arXiv:1406.3020*

Fridlund, C., 2000, in *Darwin and Astronomy: the Infrared Space Interferometer*, volume 451, p. 11

Gillessen, S., Eisenhauer, F., Perrin, G., Brandner, W., Straubmeier, C., Perraut, K., Amorim, A., Schöller, M., Araujo-Hauck, C., Bartko, H. et al., 2010, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 77340Y–77340Y

Gilmozzi, R. and Spyromilio, J., 2008, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 701219–701219

Glauber, R. J., 2006, *Reviews of modern physics*, **78**(4), 1267

Gliese, W., 1957, *Astron. Rechen-Institut, Heidelberg, 89 Seiten*, **8**, 1

Glindemann, A., Algomedo, J., Amestica, R., Ballester, P., Bauvir, B., Bugueno, E., Correia, S., Delgado, F., Delplancke, F., Derie, F. et al., 2003, in *Astronomical Telescopes and Instrumentation*, International Society for Optics and Photonics, pp. 89–100

Hanbury Brown, R., 1974, *The Intensity Interferometer* (Taylor & Francis London)

Haniff, C., 2007, *nar*, **51**, 565

Hecht 2nd, E., 1987, *Optics 2nd Edition Addison* (Wesley Publishing Company)

Helmich, F. P. and Ivison, R., 2009, *Experimental Astronomy*, **23**(1), 245

Jennison, R. C., 1958, *MNRAS*, **118**, 276

Kaltenegger, L. and Sasselov, D., 2011, *The Astrophysical Journal Letters*, **736**(2), L25

Kaltenegger, L. and Traub, W. A., 2009, *Astrophysical Journal*, **698**, 519

Kasdin, N., Vanderbei, R., Spergel, D. and Littman, M., 2003, *The Astrophysical Journal*, **582**, 1147

Kasdin, N. J., Vanderbei, R. J., Littman, M. G. and Spergel, D. N., 2005, *Applied optics*, **44**(7), 1117

Kessler, M., Steinz, J., Anderegg, M., Clavel, J., Drechsel, G., Estaria, P., Faelker, J., Riedinger, J., Robson, A., Taylor, B. et al., 1996, *Astronomy and Astrophysics*, **315**, L27

Klein, I., Guelman, M. and Lipson, S., 2007, *Applied optics*, **46**(20), 4237

Kloppenborg, B., Stencel, R., Monnier, J. D., Schaefer, G., Zhao, M., Baron, F., McAlister, H., Ten Brummelaar, T., Che, X., Farrington, C. et al., 2010, *Nature*, **464**(7290), 870

Kraus, S., Schloerb, F., Traub, W., Carleton, N., Lacasse, M., Pearlman, M., Monnier, J., Millan-Gabet, R., Berger, J.-P., Haguenauer, P. et al., 2005, *The Astronomical Journal*, **130**(1), 246

Kristiansen, R. and Nicklasson, P. J., 2009, *Acta Astronautica*, **65**(11), 1537

Labeyrie, A., 1996, *Astronomy and Astrophysics Supplement Series*, **118**, 517

Labeyrie, A., Lipson, S. G. and Nisenson, P., 2006, *An introduction to optical stellar interferometry* (Cambridge University Press)

Le Bohec, S. and Holder, J., 2006, *ApJ*, **649**, 399

LeBohec, S., Barbieri, C., de Wit, W.-J., Dravins, D., Feautrier, P., Foellmi, C., Glindemann, A., Hall, J., Holder, J., Holmes, R. et al., 2008, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 70132E–70132E

Ledrew, G., 2001, *Journal of the Royal Astronomical Society of Canada*, **95**, 32

Lim, Y.-H. and Ahn, H.-S., 2014, *Journal of the Franklin Institute*, **351**(2), 1112

Mahadevan, S., Ramsey, L., Bender, C., Terrien, R., Wright, J. T., Halverson, S., Hearty, F., Nelson, M., Burton, A., Redman, S. et al., 2012, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 84461S–84461S

Mandel, L. and Wolf, E., 1995, *Optical coherence and quantum optics* (Cambridge university press)

McAlister, H. A., Bagnuolo Jr, W. G., Theo, A., Cadman, R., Hopper, C., Ridgway, S. T., Simison, E. J., Shure, M. A., Sturmann, L. and Turner, N. H., 2000, in *Astronomical Telescopes and Instrumentation*, International Society for Optics and Photonics, pp. 465–471

Monnier, J. D., Berger, J.-P., Millan-Gabet, R. and Theo, A., 2004, in *Astronomical Telescopes and Instrumentation*, International Society for Optics and Photonics, pp. 1370–1378

Monnier, J. D., Pedretti, E., Thureau, N., Berger, J.-P., Millan-Gabet, R., ten Brummelaar, T., McAlister, H., Sturmann, J., Sturmann, L., Muirhead, P. et al., 2006, in *Astronomical Telescopes and Instrumentation*, International Society for Optics and Photonics, pp. 62681P–62681P

Monnier, J. D., Zhao, M., Pedretti, E., Thureau, N., Ireland, M., Muirhead, P., Berger, J.-P., Millan-Gabet, R., Van Belle, G., Ten Brummelaar, T. et al., 2007, *Science*, **317**(5836), 342

Murakami, H., Freund, M. M., Ganga, K., Guo, H., Hirao, T., Hiromoto, N., Kawada, M.,

Lange, A. E., Makiuti, S., Matsuhara, H. et al., 1996, *Publications of the Astronomical Society of Japan*, **48**(5), L41

Napier, P. J., Bagri, D. S., Clark, B. G., Rogers, A., Romney, J. D., Thompson, A. R. and Walker, R. C., 1994, *Proceedings of the IEEE*, **82**(5), 658

Neugebauer, G., Habing, H., Van Duinen, R., Aumann, H., Baud, B., Beichman, C., Beintema, D., Boggess, N., Clegg, P., De Jong, T. et al., 1984, *The Astrophysical Journal*, **278**, L1

Parrent Jr, G., 1960, *Optica Acta: International Journal of Optics*

Pilbratt, G., Riedinger, J., Passvogel, T., Crone, G., Doyle, D., Gageur, U., Heras, A., Jewell, C., Metcalfe, L., Ott, S. et al., 2010, *arXiv preprint arXiv:1005.5331*

Puig, L., Carpano, S., Rando, N., Stankov, A. and Escudero, I., 2010, in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics, pp. 77314I–77314I

Ridgway, S. T., van Belle, G., Mourard, D., Perrin, G., Duvert, G., Genzel, R., Haniff, C., Hummel, C., Lawson, P., Monnier, J. et al., 2011, *Proceedings of the International Astronomical Union*, **7**(T28A), 292

Schneider, J., Dedieu, C., Sidaner, P. L., Savalle, R. and Zolotukhin, I., 2011, *arXiv preprint arXiv:1106.0586*

Stepp, L. M., Daggert, L. G. and Gillett, P. E., 2003, in *Astronomical Telescopes and Instrumentation*, International Society for Optics and Photonics, pp. 309–321

Traub, W. A., 2012, *The Astrophysical Journal*, **745**(1), 20

Tschunko, H. F. A., 1983, *ao*, **22**, 133

van CITTERT, P. H., 1958, *Physica*, **24**(1), 505

Von Braun, K., Boyajian, T. S., Kane, S. R., van Belle, G. T., Ciardi, D. R., López-Morales, M., McAlister, H. A., Henry, T. J., Jao, W.-C., Riedel, A. R. et al., 2011, *The Astrophysical Journal Letters*, **729**(2), L26

Werner, M., Roellig, T., Low, F., Rieke, G., Rieke, M., Hoffmann, W., Young, E., Houck, J., Brandl, B., Fazio, G. et al., 2004, *The Astrophysical Journal Supplement Series*, **154**(1), 1