

Direct demodulation for alternative Shack-Hartmann alignment

Combining Fourier demodulation
with curvature sensing

Systems & Control Master Thesis
Seppe Kuipers

Direct demodulation for alternative Shack-Hartmann alignment

Combining Fourier demodulation
with curvature sensing

by

Seppe Kuipers

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday May 30, 2024 at 1:30 PM.

Student number: 4496248
Project duration: February 9, 2023 – May 30, 2024
Thesis committee: Prof. dr. ir. M. Verhaegen, TU Delft, supervisor
Dr. O. A. Soloviev, TU Delft, daily supervisor
Dr. R. Van de Plas, TU Delft, committee member

Cover: The VLT's new deformable secondary mirror by ESO/E. Vernet
Style: TU Delft Report Style, with modifications by Daan Zwaneveld

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Preface

I would like to express my gratitude to Oleg for supervising me during this thesis. The silly seminars were a nice constant throughout the project and helped me immensely throughout. Many problems were cleared up through these presentations.

I would also like to thank Dimitri for his feedback and his knowledge on things related to the lab.

I would like to thank my friends for letting me vent over the course of more than a year. The most important hypothesis is now confirmed: it turns out that it is actually possible to obtain a masters degree in Delft.

I would like to thank my parents, without whom of course none of this would have been possible.

And finally I would like to thank the inventor of peanut butter, Marcellus Gilmore Edson, from the bottom of my heart.

*Sepe Kuipers
Delft, May 2024*

De Tour win je in bed

JOOP ZOETEMELK

Abstract

The Shack-Hartmann wavefront sensor is a widely-used device to measure the light wavefront. Currently, the sensor is used as a gradient sensor, which is achieved by placing the microlens array in the plane conjugate to the deformable mirror and the aberration. The resulting spot pattern is then transformed into a gradient through the use of centroiding: the measurement of the relative movement of all spots compared to the reference image.

As an alternative to the traditional alignment, alternative alignment moves the Shack-Hartmann wavefront sensor: the CCD is now optically conjugated to the deformable mirror and the aberrations, and not the gradients but the curvature of the wavefront is measured. This changes the behaviour of the Shack-Hartmann sensor: instead of the amount of spots remaining the same and the CCD region occupied by them changing, the amount of spots starts to change but the CCD region occupied by them remains the same in the face of aberrations. Curvature measurements can be used directly to actuate bimorph and membrane deformable mirrors, bypassing the need for complex control schemes.

As an alternative to centroiding Fourier demodulation can also be used to convert the information measured by the wavefront sensor into gradients. For large aberrations where the spots stray out of their subapertures and the centroiding algorithm breaks down this method is especially suited.

In this thesis it is shown that a synergy exists between alternative alignment and Fourier demodulation. Large aberrations no longer occupy a huge region on the CCD, instead changing the amount of spots, and the movement of the spots outside of their subapertures is no longer a problem as there is no need for rigidly defined subapertures.

The combined use of alternative alignment and Fourier demodulation was tested on simulations, both for square and circular apertures, and for cases with extra added Poisson noise. Finally real-life experiments were also conducted using circular apertures.

Contents

Preface	i
Abstract	iii
Nomenclature	x
1 Introduction	1
1.0.1 The necessity of wavefront sensing	1
1.1 An Overview of Adaptive Optics	1
1.1.1 The Principle of Adaptive Optics	1
1.1.2 Diffraction	3
1.1.3 Image quality	4
1.1.4 Turbulence	5
1.2 Classical Gradient-based Wavefront Reconstruction	5
1.2.1 Measuring the Wavefront	6
1.2.2 Shack-Hartmann sensor	6
1.2.3 Classical wavefront reconstruction	7
1.3 Overview	10
1.4 Thesis motivation	10
2 Curvature sensing with a Shack-Hartmann sensor	11
2.1 Curvature sensing	11
2.2 Schematic comparison of alternative alignment with traditional alignment	14
2.3 Summary	15
3 Using Fourier demodulation to extract the gradients from a Hartmanngram	16
3.1 The method	16
3.2 Summary	17
4 Calculating the gradients of a known field using the finite difference	18
4.1 Calculating the finite difference from F directly	18
4.1.1 Case without aperture	18
4.1.2 Case with aperture	19
4.2 Summary	21
5 Calculating and correcting for inconsistencies of a gradient field	22
5.1 The gradient fields	22
5.2 Calculation of the correction terms	23
5.3 Summary	24
6 Extending the gradients beyond their aperture	25
6.1 The method of gradient extension	25
6.2 Summary	27
7 Integration of the gradients	28
7.1 The finite difference	28
7.1.1 Case without aperture	28
7.1.2 Case with aperture	29
7.2 Fourier integration	31
7.2.1 Discrete Fourier integration	31
7.2.2 Continuous Fourier integration	33
7.3 Summary	35

8	Using phase unwrapping to remove phase jumps	37
8.1	The method of phase unwrapping	37
8.2	Summary	41
9	Setup for the traditional and alternative alignments	42
9.1	Traditional and alternative alignment setups in the lab	42
9.2	Hartmanngram generation from the setup	45
9.3	Shack-Hartmann wavefront sensor used in the experiments	45
9.4	Obtained pictures	45
9.5	Summary	47
10	Results	48
10.1	Simulated Shack-Hartmann parameters	48
10.2	Results using a circular aperture	48
10.2.1	Geometrical optics, alternative alignment	49
10.2.2	Geometrical optics, traditional alignment	51
10.2.3	Physical optics, traditional alignment	53
10.2.4	Physical optics, alternative alignment	55
10.3	Comparison with centroiding	57
10.4	Shot noise	57
10.5	Results using a square aperture	60
10.5.1	Geometrical optics, traditional alignment	60
10.5.2	Geometrical optics, alternative alignment	62
10.6	Discussion	64
10.7	Summary	65
11	Conclusion of the thesis	66
11.0.1	Recalling the research question	66
11.1	Theoretical result	66
11.2	Results from the simulations	66
11.2.1	RMSE values for the square aperture	66
11.2.2	RMSE values for the circular aperture	67
11.2.3	Comparison with centroiding	67
11.3	Answer to the research question	68
11.4	Advice for further research	68
	References	69
A	Appendix	70

List of Figures

1.1	A general imaging system to illustrate the coordinates and planes used throughout this paper [16]	2
1.2	The concept of the wavefront illustrated [16]	3
1.3	The flattening of spherical waves with distance [4]	3
1.4	The Airy disk: the point-spread function of a diffraction-limited system [16]	4
1.5	2D schematic of a Shack-Hartmann wavefront sensor [15]	6
1.6	3D Schematic drawing showing the principle of the Shack-Hartmann wavefront sensor. [2]	6
1.7	Illustration of a) Fried, b) Hudgin, and c) Southwell sensor geometries. The x- and y-slopes are denoted by the horizontal and vertical dashes respectively, and the phase points are denoted by the dots. [13]	7
1.8	The first 6 rows of Zernike functions $Z_n^m(\rho, \theta)$ defined on a unit disk for $\rho \in [0, 1]$ and $\theta \in [0, 2\pi]$, sorted in a Zernike pyramid corresponding to $k = 0, \dots, 20$. [1]	9
2.1	Illustration of the applied control signal and the normalised density on the left, fit of error on the right. [7]	12
2.2	The standard alignment and alternative alignment shown side by side. The magnification factor for this system is equal to $M = -\frac{f_1}{f_2}$	12
2.3	Four Shack-Hartmann patterns showing the difference between the alternative and standard alignments. Notice that the total illuminated area on the CCD changes in the case of standard alignment, whereas it remains constant in the case of alternative alignment.	13
2.4	Schematic view of the difference between the traditional and alternative alignments. The wavefront propagates from top to bottom. The black bars at x' in the traditional alignment and at x in the alternative alignment denote the locations with optical conjugacy to the deformable mirror.	14
2.5	Schematic view that shows the comparison of spot patterns for the alternative alignment with the traditional alignment. Notice how the resulting spot pattern for the traditional alignment changes shape depending on the size of the aberration, while retaining the same amount of spots, whereas the alternative alignment does not stray outside the original aperture, but the amount of spots can change.	14
4.1	Example scalar field F . It is based on astigmatism.	18
4.2	The x- and y-gradient of the astigmatism aberration. Notice the dimensions of the grids are still 100x100 due to the 'dife' command.	19
4.3	Defocus aberration on a 100x100 grid. Its amplitude is equal to +0.5 at the maximum and -0.5 at the minimum. The radius of the aperture is 35 pixels. The values outside the aperture have been set to NaN as there is no information at these points.	20
4.4	The x- and y-gradient of the defocus aberration. Notice the dimensions of the grids are still 100x100 due to the 'dife' command, but the shape of the circles has been changed slightly due to the NaN values.	21
6.1	Gradients retrieved from a defocus aberration. The white regions are set to NaN values as there is no information from the Shack-Hartmann sensor.	26
6.2	Extended gradients retrieved from a defocus aberration. The remaining NaN values after the extension operation have been replaced with zeroes.	26
6.3	Extended and seamed gradients retrieved from a defocus aberration. The rows in x-direction and the columns in y-direction all sum up to zero.	27
7.1	The x- and y-gradient of the astigmatism aberration. Notice the dimensions of the grids are still 100x100 due to the 'dife' command.	29

7.2	The reconstruction of the field and its reconstruction error. Notice how in the case of a full aperture the reconstruction is near perfect.	29
7.3	Ground truth field for integration.	30
7.4	Gradients from the finite difference.	30
7.5	The extended x- and y-gradients.	31
7.6	The reconstruction and its error using line integration in the case of an aperture.	31
7.7	A visual representation of the way the derivative $S_x(n + \frac{1}{2}, m) \equiv \Phi(n + 1, m) - \Phi(n, m)$ is taken in one dimension. Δ represents the step size, for real systems Δ represents the distance between the centers of two pixels on the CCD.	32
7.8	Fourier transforms of gradients obtained by the finite difference.	32
7.9	The Fourier transform of the ground truth, the reconstruction, and the difference between the two.	33
7.10	The ground truth, the reconstruction, and their differences using the discrete Fourier kernel.	33
7.11	Seamed and extended Linear x- and y-gradients retrieved from a Shack-Hartmann wavefront sensor.	34
7.12	Fourier transforms of the seamed and extended Linear x- and y-gradients retrieved from a Shack-Hartmann wavefront sensor.	34
7.13	Fourier transform of the wavefront calculated using the continuous Fourier kernel.	35
7.14	The ground truth, the reconstruction, and their differences using the continuous Fourier kernel.	35
8.1	An illustration of what happens when the residues are not corrected for. Long lines propagate throughout the solution.	39
8.2	An illustration of what happens when branch cuts are put in place to correct for the residues. The residues will not propagate throughout the solution. Compare with figure 8.1 to see the improvement.	40
8.3	An illustration of what happens when the data is noisy to the extent that a large number of branch cuts have to be made: the resulting reconstruction is also mostly corrupted.	41
9.1	Setup of the system. A point source laser with a wavelength of 635 nm emits light that is collimated by a lens. A pinhole regulates the beam diameter of the system. Two mirrors are used to ensure the beam is shot straight at the Shack-Hartmann sensor. Depending on the position of the Shack-Hartmann sensor, the aberration lens is placed either in the plane optically conjugate to the microlens array or the charge coupled device. The Shack-Hartmann sensor is placed on a moving stage in order to switch between the standard and alternative alignment modes.	44
9.2	Images obtained from the setup in alternative alignment. From top left clockwise the flat reference image, the 50 mm lens, the 75 mm lens, the 150 mm lens and the 200 mm lens.	46
9.3	Images obtained from the setup in standard alignment. From top left clockwise the flat reference image, the 50 mm lens, the 75 mm lens, the 150 mm lens and the 200 mm lens.	46
10.13	Defocus aberration Hartmanngram with shot noise.	58
10.14	The resulting gradients in the case of shot noise on the Hartmanngram. Notice the noise is quite severe, these are gradients of a defocus aberration, but they are no longer nicely linear.	58
10.15	The ground truth, the reconstruction, and their absolute difference for the case of added shot noise on the Hartmanngram. Notice that despite the gradients being rather noisy, the algorithm can still reconstruct a wavefront that somewhat accurately describes the original.	59
10.16	Alternative aligned Shack-Hartmanngram showing a defocus aberration with shot noise.	59
10.17	The resulting gradients in the case of shot noise on the Hartmanngram in the alternative alignment. Notice the noise is quite severe, these are gradients of a defocus aberration, but they are no longer nicely linear.	60

10.18	The ground truth, the reconstruction, and their absolute difference for the case of added shot noise on the Hartmanngram. Notice that despite the gradients being rather noisy, the algorithm can still reconstruct a wavefront that somewhat accurately describes the original.	60
-------	--	----

List of Tables

8.1	An inconsistent phase field. All numbers are fractions of a full cycle.	37
8.2	The inconsistency in the field. As no jumps of more than half a cycle should be possible, the jump from 0.8 to 0.0 should be seen as a +0.2 jump from 0.8 to 1.0 instead of a -0.8 jump from 0.8 to 0.0. Summing clockwise results in a residue of +1.	38
8.3	The result of summing all clusters of four points reveals there is only one residue: a positive one in the center. This means there is a net nonzero residue present in the system.	38
8.4	An illustration of adjacent points differing more than half a cycle in inconsistent phase field. The two points in red differ more than half a phase, just like the points in orange.	38
8.5	An illustration of the futility of trying to get rid of inconsistencies in a field by adding or removing one phase. The location of the inconsistency changes (depending even on the direction of correction applied) but the inconsistencies themselves remain.	38
10.1	Table showing the focal length and pitch used in the two different scenarios.	48
10.2	RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.	57
10.3	Table showcasing the difference in reconstruction coefficients for the Fourier demodulation method versus the centroiding method. Notice how the Fourier reconstruction coefficient is closer to the ground truth coefficient than the FrontSurfer coefficient, and how the residue is small for both circumstances.	57
10.4	The different values of the RMSE shown in a table. Not that for positive defocus in a square aperture the traditional alignment actually outperforms the alternative alignment.	64
11.1	RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.	67
11.2	RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.	67
11.3	Table showcasing the difference in reconstruction coefficients for the Fourier demodulation method versus the centroiding method. Notice how the Fourier reconstruction coefficient is closer to the ground truth coefficient than the FrontSurfer coefficient, and how the residue is small for both circumstances.	68

Nomenclature

Abbreviations

Abbreviation	Definition
AO	Adaptive optics
CCD	Charge-coupled device
DM	Deformable mirror
HRI	High-resolution imaging
MLA	Microlens array
RMS	Root mean square
RMSE	Root mean square error
SH	Shack-Hartmann
SHWFS	Shack-Hartmann wavefront sensor
WFS	Wavefront sensor

1

Introduction

1.0.1. The necessity of wavefront sensing

Wavefront measurements play a crucial role in today's society. All large telescopes in the world depend on wavefront sensing and correction using a deformable mirror. Wavefront sensing is used in the industry to detect imperfections in manufacturing, and the medical world uses wavefront sensing to determine the aberration in a patient's eye. Wavefront sensing is an important and widely used tool, and as such, a lot of effort is spent on coming up with ways to improve the performance of currently existing wavefront sensing methods. This thesis aims to contribute to these improvements.

1.1. An Overview of Adaptive Optics

The basic principle for adaptive optics throughout this thesis is as follows: a wavefront sensor measures the incoming wavefront for deformations. This information is used to calculate which deformations are present, and how to counteract them using a deformable mirror. These calculated corrections are then sent to the deformable mirror, which changes shape to make the incoming wavefront as flat as possible. This main focus of this thesis lies in exploring the combination of Erez N. Ribak's method of Fourier demodulation as a way to read out a Shack-Hartmann sensor with Oleg A. Soloviev's alternative physical setup in which the CCD of the wavefront sensor is placed in the plane optically conjugated to the deformable mirror. In the subsequent parts of the introduction the following things will be dealt with:

- The principle of adaptive optics
- Shack-Hartmann sensor and wavefront reconstruction method
- Reconstruction methods

1.1.1. The Principle of Adaptive Optics

To start, the conventions and basic principles and ideas used throughout the paper are laid out. The object is located in the object plane which has coordinates $\boldsymbol{\xi} = (\xi, \eta)$, the pupil is located in the pupil plane which has coordinates $\mathbf{x} = (x, y)$, and the image is located in the image plane which has coordinates $\mathbf{u} = (u, v)$. The line going from the object to the image is called the optical axis z . The object plane and the image plane are said to be optically conjugate planes. This convention can be seen in figure 1.1.

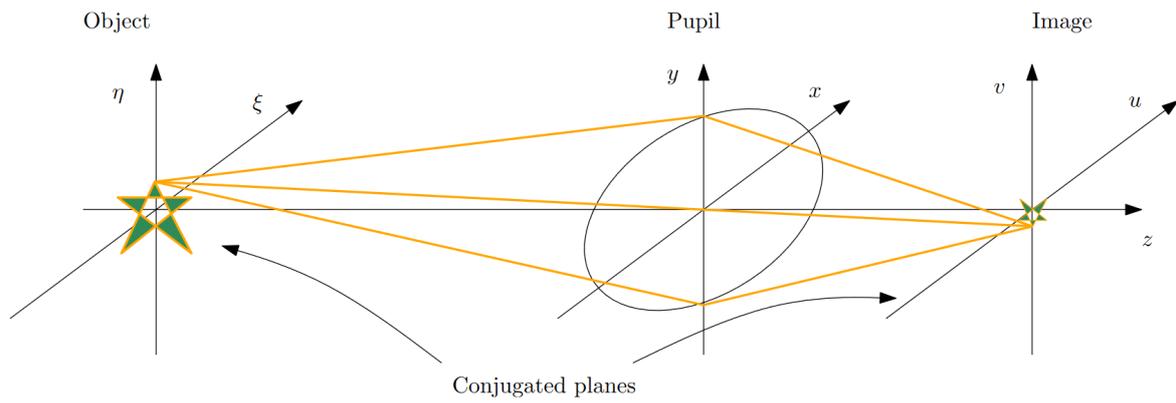


Figure 1.1: A general imaging system to illustrate the coordinates and planes used throughout this paper [16]

Light is emitted in waves that travel across planes in \mathbb{R}^2 through time t . The waves emitted at the surface of the object are defined by

$$w(\mathbf{x}) = a(\mathbf{x}) \exp(-i\phi(\mathbf{x})) \quad (1.1)$$

where a is the amplitude of the wave, and ϕ its phase. Light emitted at the same time instance t at the surface of the object is called a wavefront. The definition of a wavefront is that its phase ϕ is constant across its surface. In figure 1.2 multiple wavefronts can be seen.

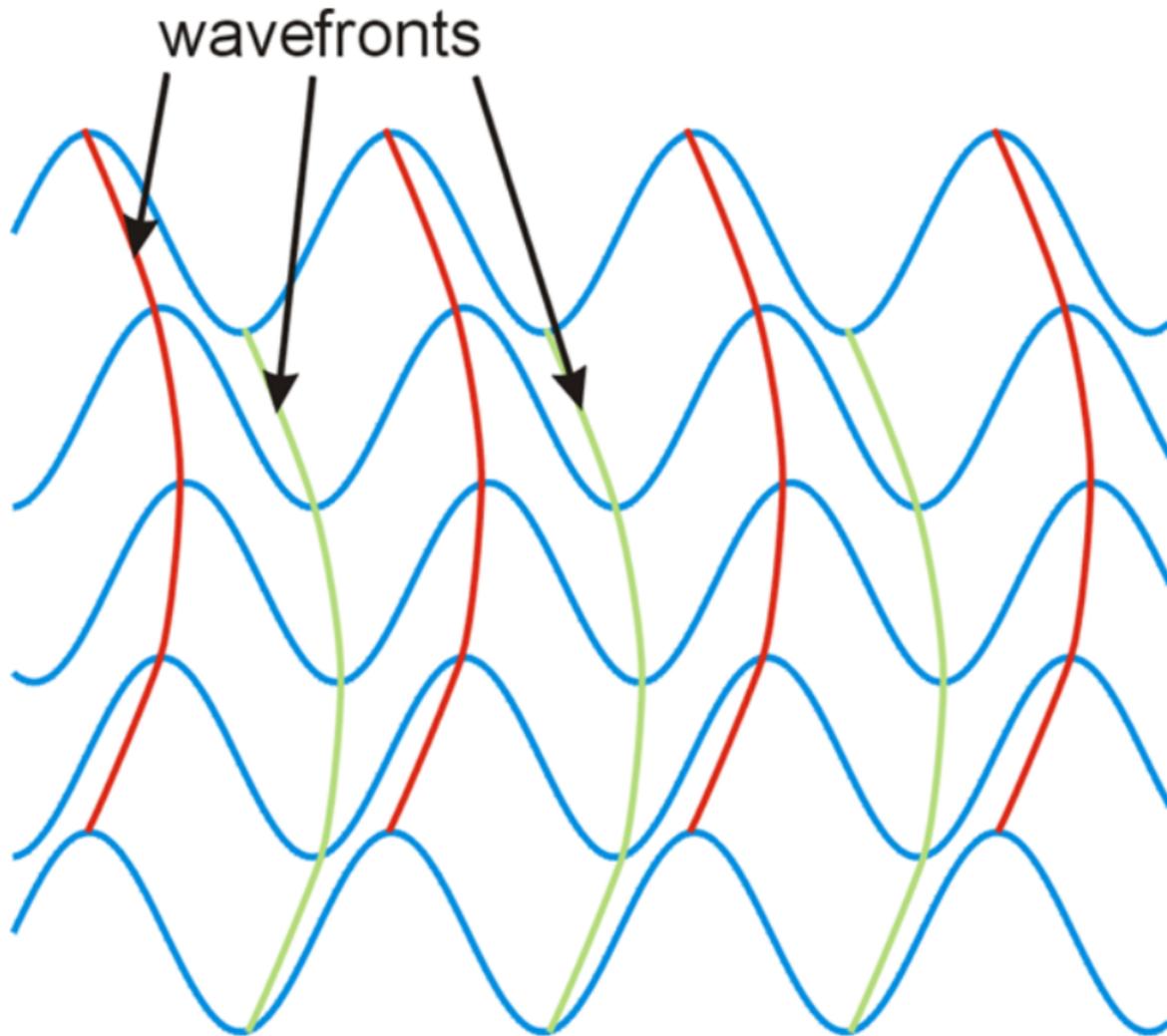


Figure 1.2: The concept of the wavefront illustrated [16]

As these waves set out, their shape is spherical, and as they travel further and further their shape slowly changes into a more flat one, until after an infinite amount of distance the resulting wave is perfectly flat. This phenomenon is illustrated in figure 1.3.



Figure 1.3: The flattening of spherical waves with distance [4]

1.1.2. Diffraction

Because of the difference in shape, physicists have divided the space behind an aperture into two regions. One is the near-field region, the region where light waves are still mostly spherical. Here the *Fresnel diffraction* or *near-field diffraction* is used to calculate the propagation of waves through the field. It is

equal to: [3]

$$A(x_o, y_o) \approx \frac{e^{-ik\Delta z}}{i\lambda\Delta z} \iint_{\mathbb{R}^2} A(x, y) e^{-i\frac{k}{2\Delta z}((x_o-x)^2+(y_o-y)^2)} dx dy \quad (1.2)$$

For the far-field, the region where light waves lose their spherical shape and become mostly flat, *Fraunhofer diffraction*, or *far-field diffraction* is used to calculate the propagation of light waves through the field. To define the transition from near-field to far-field the dimensionless *Fresnel number*, defined as $F = r^2/(L \cdot \lambda)$, is used, where r is the radius of the aperture, L is the distance from the aperture, and λ is the wavelength of light. If $F > 1$ the near-field equation has to be used, and if $F \leq 1$ the Fraunhofer diffraction approximation can be used. The *Fraunhofer diffraction integral* is equal to [3]

$$A(x_o, y_o) \approx \frac{e^{-ik\Delta z} e^{-i\frac{k}{2\Delta z}(x_o^2+y_o^2)}}{i\lambda\Delta z} \iint_{\mathbb{R}^2} A(x, y) e^{-i\frac{k}{2\Delta z}(x_o x + y_o y)} dx dy \quad (1.3)$$

Diffraction is a process by which the light is spread out as a result of passing through an aperture or across an edge. The maximum image quality that can therefore be achieved using adaptive optics is to make the system diffraction-limited. This means that the residual error of the wavefront reconstruction should be smaller than the effects of diffraction.

1.1.3. Image quality

The Fraunhofer diffraction approximation is equal to taking the Fourier transform of the input field $A(x)$ multiplied by the pupil function $P(x)$ – equal to one inside the systems aperture, and zero outside of it – and a phase term. Looking only at the intensity, as in a Shack-Hartmann sensor, the squared field distribution is given by [3]

$$I(\mathbf{u}) = \frac{1}{\lambda^2 f^2} \left| \iint_{\mathbb{R}^2} A(\mathbf{x}) P(\mathbf{x}) e^{-i\frac{2\pi}{\lambda f}(xu+yv)} d\mathbf{x} \right|^2 \equiv \frac{1}{\lambda^2 f^2} |\mathcal{F}(A(\mathbf{x})P(\mathbf{x}))|^2 \quad (1.4)$$

The point spread function (PSF) is the image of a point source. The PSF is the impulse response of the optical system. It is equal to the image of a point p_0 that is affected by diffraction and possible aberrations. The resulting image p_1 contains information on the aberrations present in the system. If the system is free of aberrations i.e. diffraction-limited, meaning $\phi(\mathbf{x})$ is a constant, its PSF is equal to the *Airy pattern* or *Airy disk*.

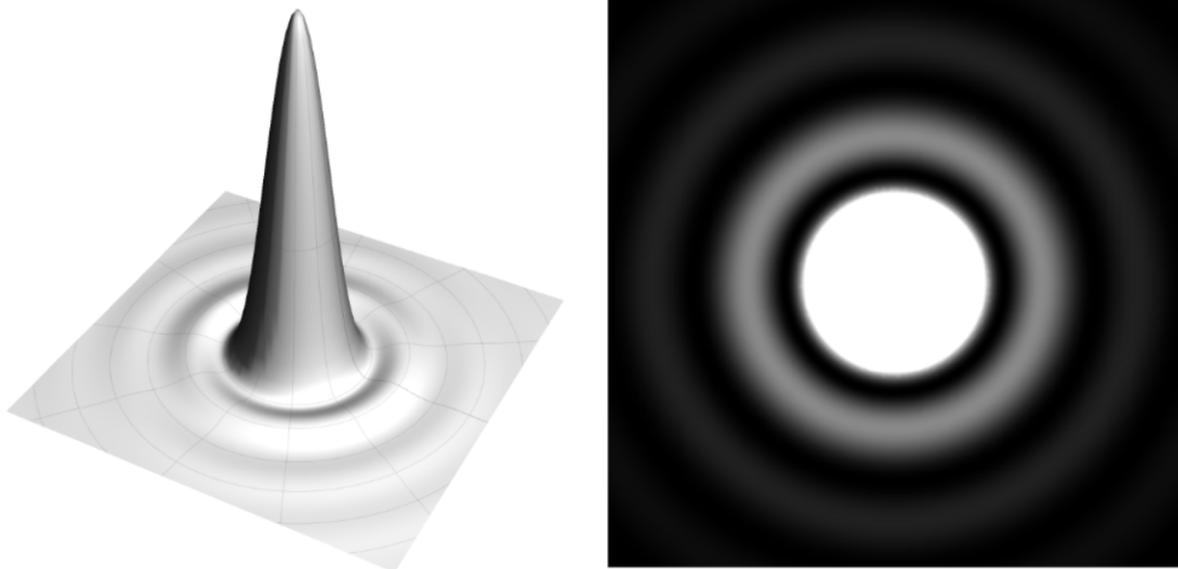


Figure 1.4: The Airy disk: the point-spread function of a diffraction-limited system [16]

On the left side of figure 1.4 is a 3-D plot of the Airy function, and on the right side a top-view is seen. The inner black ring defines the diffraction-limited resolution, and is located at a distance $1.22\lambda/D$

from the center. For systems that are not diffraction-limited, the ring contracts and its distance to the center shrinks. The PSF is equal to [3]

$$p(\mathbf{u}) = \left| \mathcal{F} \left(P(\mathbf{x}) e^{-i\phi(\mathbf{x})} \right) \right|^2 \quad (1.5)$$

The distance from the center of the PSF to the first dark ring can be used to assess the quality of an image. The angular distance $1.22\lambda/D$ mentioned above is *Rayleigh's resolution criterion*. More resolution criteria exist: *Abbe's resolution criterion* and *Sparrow's resolution criterion*. These are defined as a distance of λ/D and $0.94\lambda/D$ respectively. For Rayleigh's resolution criterion the minimum resolvable separation between two points is the diameter of the central disk of the PSF. Out of the three, Rayleigh's is strictest. Abbe's resolution criterion is more lenient and is used in the world of microscopy. Sparrow's resolution criterion is the most lenient of all, being the ultimate minimal distance needed before the two patterns form a single new one.

1.1.4. Turbulence

In astronomy, atmospheric turbulence is the primary source of aberration which decreases the resolution of the optical system. Turbulence can be described using Kolmogorov's statistical theory [5]. An important parameter in this model is the *Fried parameter* or the *Fried coherence length*, denoted by r_0 . It describes the quality of optical transmission through the atmosphere due to turbulence. To be precise, it describes the diameter of a circle over which the root means square wavefront aberration due to turbulence is equal to 1 radian. Typical values lie between 5-20 cm, with 5 cm being only average seeing conditions and 20 cm being about the best you can hope for on earth. For images with long exposure time the Fried coherence length is the upper limit at which increasing the diameter of the aperture becomes pointless. This means that without adaptive optics it would be pointless to construct telescope mirrors of more than 20 cm in diameter, which would severely limit the amount of objects that can be seen, as Rayleigh's resolution criterion states that the diameter of the aperture needs to be as large as possible to achieve high quality images.

To illustrate how an adaptive optics feedback control system would operate, an example is given where a distant star is imaged. The star, located at a large distance from the optical system, emits light which travels to the earth, where it arrives with its wavefront still flat, as the vacuum of space does not cause any aberrations. The light then passes through the atmosphere where temperature, wind speed, and pressure vary greatly between its different layers. This turbulence causes the wavefront to distort. Once it reaches the optical system located on earth's surface the aberrations have to be removed to get a coherent image of the star. The light is captured by a telescope system and reflected off a deformable mirror. The deformable mirror then directs the beam onto a beam splitter that divides the light between the science camera and the wavefront sensor. The wavefront sensor analyses the disturbance of the wavefront and sends a control input to the deformable mirror to change its shape, counteracting the disturbance. To do this it needs to copy the negative shape of the measured deformation. In the ideal scenario the system is fast and precise enough to make the system diffraction-limited, meaning the outer telescope lens is the inhibiting factor of the image quality. Adaptive optics plays a huge role in the telescoping world, as exposure times can be as long as 12 hours to compensate for the low amount of distant light reaching the earth. Without the wavefront sensor analysing the aberrations and the deformable mirror correcting them, turbulence would render any image with long exposure times useless.

1.2. Classical Gradient-based Wavefront Reconstruction

This subsection details how to reconstruct the wavefront using a Shack-Hartmann wavefront sensor in combination with either modal or zonal reconstruction algorithms. In the first subsection, an overview is given of the Shack-Hartmann wavefront sensor. The modal and zonal algorithm approaches are explained in the second subsection.

1.2.1. Measuring the Wavefront

A range of sensors exist to measure the wavefront. Two important classes can be distinguished: the *pupil plane* sensors and the *focal plane* sensors. Pupil plane sensors are placed in the plane optically conjugated to the system pupil, whereas focal plane sensors measure the wavefront in their focal plane. This paper deals only with pupil plane sensors and will elaborate on two of them: the *Shack-Hartmann wavefront sensor* and the *curvature sensor*.

1.2.2. Shack-Hartmann sensor

The Shack-Hartmann sensor is a widely used tool to analyse the wavefront. It is a modified version of the Hartmann sensor. The original Hartmann sensor from 1904 operates using an array of tiny holes through which the incoming light passes. The light then goes on to hit a photoreceptive plate located behind the grid. The aberrations in the wavefront are then measured by approximating them as a network of piecewise tip/tilt aberrations. In 1971 Shack and Platt [8] modified the Hartmann sensor by replacing the holes with lenslets. This greatly improved the light gathering ability or photon efficiency of the sensor. This sensor has greatly surpassed the original Hartmann sensor in popularity, though it should be noted that the original Hartmann sensor is still used today, mainly in the field of (extreme) ultraviolet microscopy and lithography, for the simple reason that ultraviolet light is blocked by the glass of the lenslets.

The Shack-Hartmann wavefront sensor is a popular pupil plane sensor. It operates by sampling the wavefront across a grid of lenslets. This method has the advantage of approximating the wavefront piecewise linearly, meaning it is fast enough for real-time control, but the downside is that the total wavefront will be a piecewise linear approximation of the wavefront, which is nonlinear. As the wavefront reconstruction speed plays an important role in adaptive optics, this approximation is acceptable. Below are two figures showing schematically the way a Shack-Hartmann sensor operates.

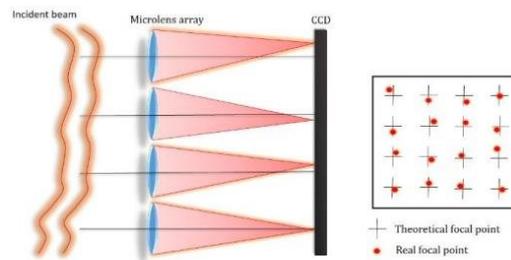


Figure 1.5: 2D schematic of a Shack-Hartmann wavefront sensor [15]

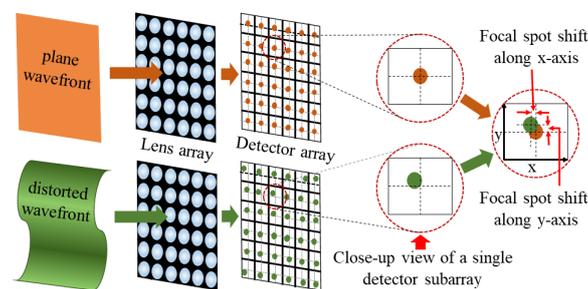


Figure 1.6: 3D Schematic drawing showing the principle of the Shack-Hartmann wavefront sensor. [2]

Figure 1.5 shows a 2D-view of the Shack-Hartmann sensor. The incoming distorted wavefront is sampled by the microlens array, which focuses the light on the CCD located behind it. In Figure 1.6 the same is seen 3D. The shifting of the spots provides measurements for Δx and Δy , from which the wavefront

can be reconstructed. For each subaperture located in the i -th row and j -th column in the lenslet array, the local displacements are calculated using:

$$\begin{aligned}\Delta x(i, j) &= \kappa_x \frac{\partial \phi(x_i, y_j)}{\partial x} + \eta_x(i, j) \\ \Delta y(i, j) &= \kappa_y \frac{\partial \phi(x_i, y_j)}{\partial y} + \eta_y(i, j)\end{aligned}\quad (1.6)$$

where the η 's represent the measurement noise and the effect of higher order aberrations that can not be modelled by a simple tilt, and the κ 's are constants determined by the optical parameters such as pupil size, distance between pupil and lens, etc.

The incoming light in a Shack-Hartmann sensor is often scattered across multiple pixels on the CCD chip because of diffraction. This means the location of the center can be calculated using the following equations:

$$\begin{aligned}s_x(i, j) &= \gamma_x \frac{\sum_{u,v} u \Delta_x I(u, v)}{\sum_{u,v} I(u, v)} \cong \frac{\partial \phi(x_i, y_i)}{\partial x} + \tilde{n}_x(i, j) \\ s_y(i, j) &= \gamma_y \frac{\sum_{u,v} v \Delta_y I(u, v)}{\sum_{u,v} I(u, v)} \cong \alpha_y \frac{\partial \phi(x_i, y_i)}{\partial y} + \tilde{n}_y(i, j),\end{aligned}\quad (1.7)$$

where $I(u, v)$ is equal to the intensity measured by the pixel in the u -th row and v -th column of the camera, Δ_x and Δ_y equal to the spacing of the pixels along the x- and y-axis, γ_x and γ_y are approximately $\frac{1}{z}$, where z is the propagation distance between the aperture and detection planes.

1.2.3. Classical wavefront reconstruction

The classical approaches to wavefront analysis can be divided in two camps: the zonal (local) and modal (global) methods. The zonal methods work by dividing the sensor area into a rectangular grid of points. Then, using finite differences, the aberration is calculated in all of the grid points. Different grid geometries exist. These will be presented in the next subsection. The modal approach operates using a set of basis functions which are matched to the measured wavefront. Popular basis functions are the Zernike polynomials and the Karhunen-Loève functions. The use of basis functions is preferred over the use of zonal reconstruction when it comes to error propagation, ease of use, and speed [14].

Finite Difference Methods

The most popular zonal method is the finite difference method. It subdivides the sensor area into a rectangular grid of points. Different zonal algorithms exist; popular ones include the Fried, Hudgin, and Southwell geometries. These geometries subdivide the grid into subapertures, and within each of these subapertures the local wavefront is then calculated using a least-squares approach. The total wavefront is then obtained by combining all the subapertures. Below is a figure illustrating different geometries

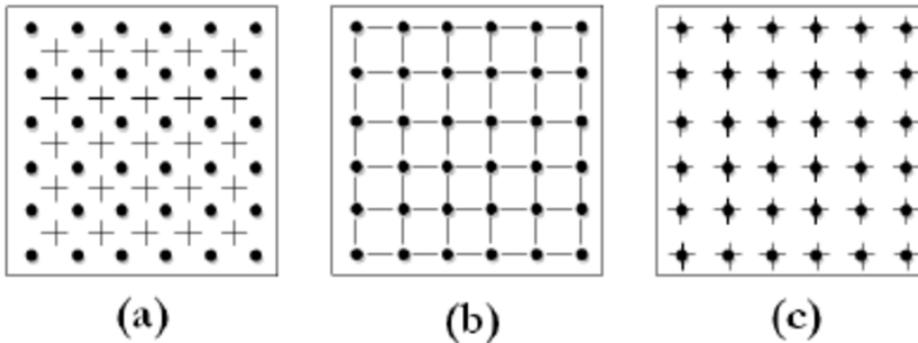


Figure 1.7: Illustration of a) Fried, b) Hudgin, and c) Southwell sensor geometries. The x- and y-slopes are denoted by the horizontal and vertical dashes respectively, and the phase points are denoted by the dots. [13]

For Fried's finite difference model the slopes are approximated in the following way:

$$\frac{\partial\phi(x_i, y_j)}{\partial x} \approx [(\phi(x_{i+1}, y_{j+1}) + \phi(x_{i+1}, y_j)) - (\phi(x_i, y_j) + \phi(x_i, y_{j+1}))] / (2D_L)$$

where D_L is equal to

$$\frac{\partial\phi(x_i, y_j)}{\partial y} \approx [(\phi(x_{i+1}, y_{j+1}) + \phi(x_i, y_{j+1})) - (\phi(x_i, y_j) + \phi(x_{i+1}, y_j))] / (2D_L),$$

the size of the subaperture and $i, j = 1, \dots, n$. Now, using the centroid algorithm formulae

$$s_x(i, j) = \gamma_x \frac{\sum_{u,v} u \Delta_x I(u, v)}{\sum_{u,v} I(u, v)} = \alpha_x \frac{\partial\phi(x_i, y_i)}{\partial x} + \tilde{n}_x(i, j)$$

$$s_y(i, j) = \gamma_y \frac{\sum_{u,v} v \Delta_y I(u, v)}{\sum_{u,v} I(u, v)} = \alpha_y \frac{\partial\phi(x_i, y_i)}{\partial y} + \tilde{n}_y(i, j),$$

The slopes can be calculated using

$$\begin{bmatrix} s_x(i, j) \\ s_y(i, j) \end{bmatrix} - \frac{\alpha}{2D_L} \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \phi(x_i, y_j) \\ \phi(x_i, y_{j+1}) \\ \phi(x_{i+1}, y_j) \\ \phi(x_{i+1}, y_{j+1}) \end{bmatrix} + \begin{bmatrix} n_x(i, j) \\ n_y(i, j) \end{bmatrix},$$

where n_x and n_y are noise terms. One can now proceed by stacking the spatial measurements $\begin{bmatrix} s_x(i, j) \\ s_y(i, j) \end{bmatrix}$ for all $i, j = 1, \dots, n$ into one big vector s . This vector can then be used to calculate the global field using the least squares approach:

$$s = \mathbf{G}\phi + n,$$

Where G is a matrix which depends on the chosen geometry of the approximation of the spatial derivatives. As can be seen, there are at every i, j fewer equations than unknowns. This means that statistical information on the variance of the wavefront is necessary:

$$E[\phi\phi^T] = C_\phi.$$

This means the problem can be written into the following form:

$$\begin{aligned} & \text{minimize} && \phi\epsilon^T\epsilon \\ & \text{subject to} && s = G\phi + L_n\epsilon \end{aligned}$$

The solution of this weighted least squares problem is given by

$$\tilde{\phi} = C_\phi G^T (GC_\phi G^T + C_n)^{-1} s$$

Modal Methods

Instead of approximating the global wavefront by interpolating using a rectangular grid, a linear combination of base functions can be used to fit the measured aberrations. Popular base functions include the Zernike polynomials and the Karhunen-Loève functions. Zernike functions are orthogonal to one another and always reach a maximum magnitude of 1 at the aperture boundary. The Zernike functions can be arranged in a *Zernike pyramid*. The first 21 Zernike functions, ordered in the pyramid, look like this:

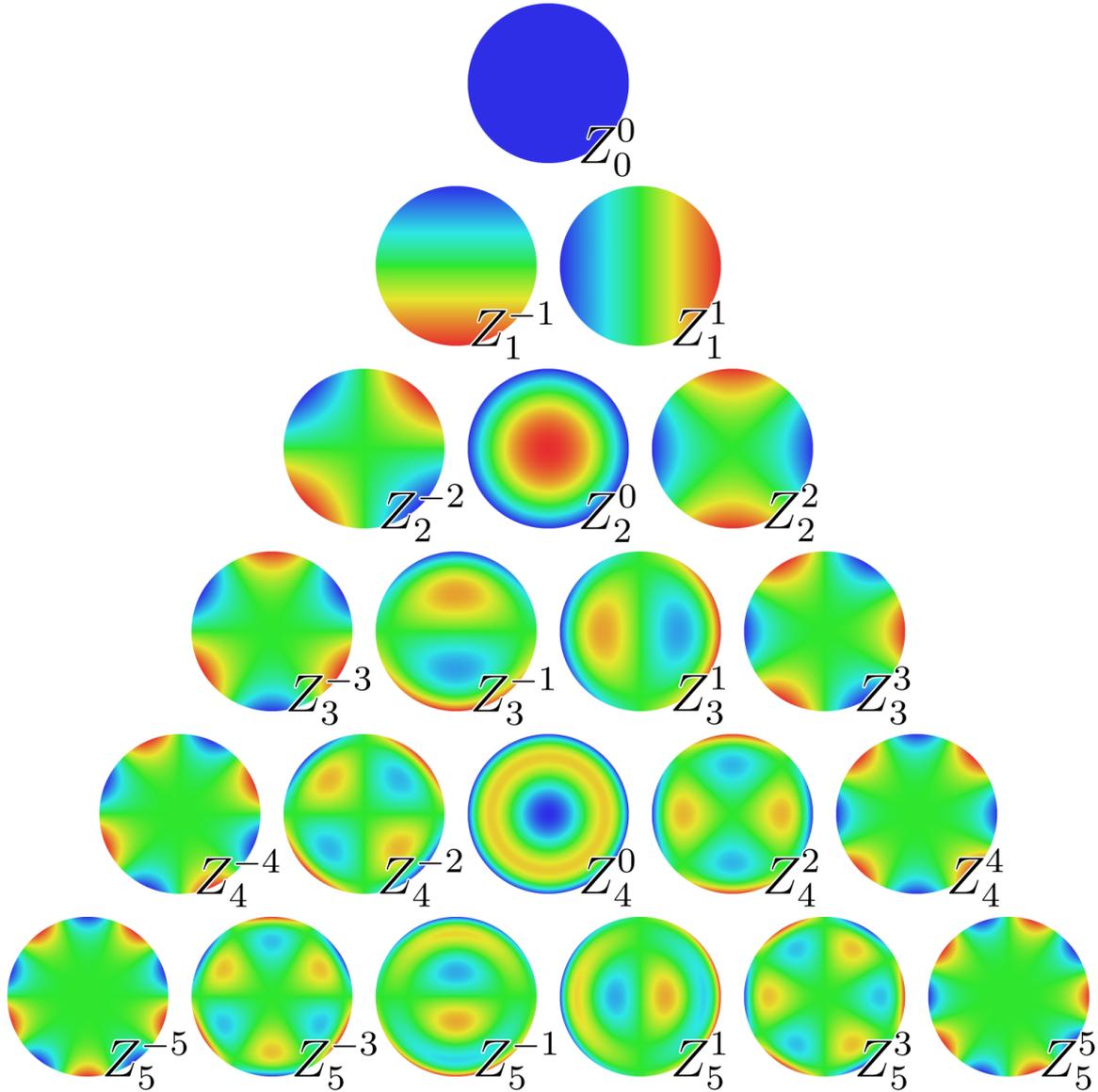


Figure 1.8: The first 6 rows of Zernike functions $Z_n^m(\rho, \theta)$ defined on a unit disk for $\rho \in [0, 1]$ and $\theta \in [0, 2\pi]$, sorted in a Zernike pyramid corresponding to $k = 0, \dots, 20$. [1]

The measured wavefront can be approximated via

$$\phi(x, y) = \sum_{k=1}^K z_k Z_k(x, y), \quad (1.8)$$

where K denotes the highest order of polynomial used, z_k the weighting coefficient, and Z_k the particular aberration, for example tip/tilt, defocus, astigmatism etc. The small index k is the Noll-index. It counts through the pyramid from left to right and top to bottom: Z_0^0 in the figure corresponds to $k = 0$, Z_1^{-1} and Z_1^1 correspond to $k = 1$ and $k = 2$ respectively and so forth. It should be noted that the piston term ($k = 0$) is left out, as it is a constant offset, meaning it is invisible to gradient-based sensors.

The wavefront aberrations can now be calculated by equating the slopes to the sum of basis functions

and a noise term:

$$\begin{aligned} s_x(i, j) &= \sum_{k=1}^K z_k \frac{\partial Z_k(x_i, y_i)}{\partial x} + \eta_x(i, j) \\ s_y(i, j) &= \sum_{k=1}^K z_k \frac{\partial Z_k(x_i, y_i)}{\partial y} + \eta_y(i, j), \end{aligned} \quad (1.9)$$

where K denotes the highest order polynomial, (x_i, y_i) denotes the location of the lenslet in the i -th row and j -th column. It can be written in matrix notation

$$s = \mathbf{B}\zeta + \eta, \quad (1.10)$$

where s is a vector containing all slopes in x- and y-direction: $s = [s_x^T, s_y^T]^T$, \mathbf{B} is a matrix containing partial derivatives of K Zernike functions, and ζ is the unknown: the coefficients z_k describing the size of each of the individual Zernike functions. η is an added noise term again. A least-squares approach is used to arrive at an optimal solution. It is defined as

$$\hat{\phi} = \mathbf{B}^+ s, \quad (1.11)$$

where \mathbf{B}^+ denotes the Moore-Penrose inverse or pseudo-inverse, which exists even for singular matrices. Using this method it is possible to extract the piston term by looking at the average of the normalised basis functions. The average of their sizes will reveal the amount of piston.

1.3. Overview

For some high resolution imaging systems it is necessary to incorporate a deformable mirror with adaptive optics. This chapter showed the general principle of adaptive optics. This is the basis on which the rest of the thesis is built. The next chapter will deal with classical wavefront reconstruction using a gradient-based method. In the introduction the general principles of adaptive optics were shown. In this chapter the standard method of centroiding using both a classically aligned SH sensor is shown. Two methods for calculating the wavefront were presented: the finite difference and the modal method. These are the conventional methods of wavefront reconstruction, and they will serve as a benchmark for the alternative alignment and Fourier demodulation methods to be tested against. In the next chapter the Fourier demodulation method will be shown.

1.4. Thesis motivation

This thesis will investigate the combined use of the alternative alignment of the Shack-Hartmann sensor and the method of Fourier demodulation to process the Hartmanngrams resulting from the alternative alignment.

The alternative alignment turns the Shack-Hartmann sensor from a gradient sensor into a curvature sensor by establishing optical conjugacy between the CCD and the aberration. The alternative alignment is able to handle large amplitude aberrations as the region on the CCD does not shrink or expand. This has its limits of course, adding or removing too many spots will eventually break down the measurement, but the alternative alignment should outperform the traditional alignment.

The Fourier demodulation method is also a good method to use when dealing with large amplitude aberrations: whereas the centroiding algorithm needs the spots to stay within their subapertures and can only tolerate so many missing spots, the Fourier algorithm can easily deal with large spot movements and the addition or removal of spots. The research question to be answered is as follows. Can the alternative alignment combined with Fourier demodulation be used to reconstruct the wavefront? If so, which Shack-Hartmann properties are beneficial to this combination, and which types of aberrations are beneficial?

2

Curvature sensing with a Shack-Hartmann sensor

The Shack-Hartmann wavefront sensor is used to measure derivatives of the wavefront. Different derivatives can be measured. In the introduction it was seen that the traditional alignment of the Shack-Hartmann sensor results in the first order derivatives or gradients of the wavefront in two perpendicular directions: $\Delta_x\Phi$ and $\Delta_y\Phi$. By aligned the Shack-Hartmann sensor differently it is possible to measure the curvature of the wavefront: $\rho(x, y) = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = \nabla^2\Phi$. This chapter is devoted to showing how this is achieved and what its consequences are.

2.1. Curvature sensing

To measure the curvature of the wavefront Vdovin, Verhaegen, and Soloviev [7] propose using a single Shack-Hartmann sensor with a standard microlens array, positioned in such a way that the CCD is optically conjugated to the deformable mirror and system aperture. This is easier than earlier methods of curvature sensing which had to measure the intensity in two planes [12], use an astigmatic MLA, or use three SH sensors.

The alternative alignment establishes a direct geometric correspondence between the coordinates on the deformable mirror surface and the sensor chip. The change in the local spot density corresponds to the Laplacian curvature of the mirror. In the traditional setup, the deformable mirror and the microlens array of the wavefront sensor are conjugated to the system pupil, see figure 2.2, which means every subaperture corresponds to a particular pupil subaperture and to the corresponding patch of the deformable mirror. The beam crosses the microlens array in a fixed area and the spot pattern can move over the camera chip but contains the same number of spots under certain obvious conditions.

In the alternative alignment the focal plane of the microlens array, the surface of the CCD, is located in the plane conjugated to the pupil as seen in figure 2.2. In this case a direct geometric correspondence exists between the coordinates on the deformable mirror surface and the sensor chip. The microlens array is now located in front of the aperture and the beam can cross it in different places depending on the wavefront shape. The number of spots in the Hartmann pattern is no longer fixed. Instead, the region they occupy becomes fixed. The change in local density of the spots is directly proportional to the local Laplacian curvature, and shows almost no dependence on intensity variations. As a consequence the control signal applied to an actuator is proportional to the integral of the points density over its area. The boundary conditions along the edge of the pupil are given by the centroids displacements along the pupil edge.

An analysis is made of the total amount of spots for three scenarios: one where there is no control input to the deformable mirror, one where there is a positive maximum signal to all elements of the deformable mirror, and one where there is a negative maximum signal to all the elements. This will establish the range of spots over which the system can operate. Different actuators are matched to different portions of the region, and the more spots are present in their region compared to the situation where there is no control input, the higher their control input. The method then uses the four nearest connected pixels

to model the spot, as it is a centroiding-based algorithm.

The density of the spots was now measured for 21 control inputs of $-1, -0.9, \dots, 0.9, 1$. The resulting normalised density displays good linear dependence on the control input, see figure 2.1.

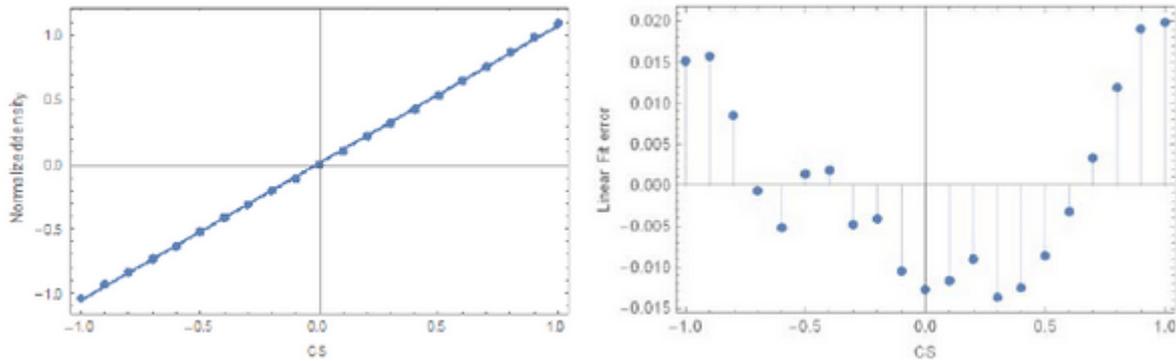


Figure 2.1: Illustration of the applied control signal and the normalised density on the left, fit of error on the right. [7]

The difference between the standard and alternative alignment of the Shack-Hartmann sensor can be seen below.

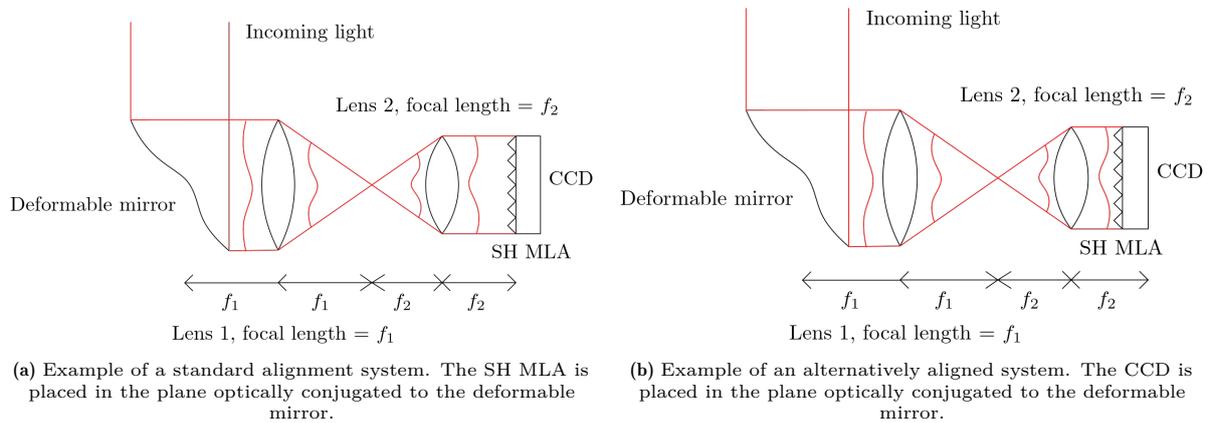


Figure 2.2: The standard alignment and alternative alignment shown side by side. The magnification factor for this system is equal to $M = -\frac{f_1}{f_2}$.

As can be seen in figure 2.2 plane optically conjugated to the deformable mirror is occupied either by the microlens array in the standard alignment or the CCD in the alternative alignment. The difference in aberrations measured between the two methods is shown below:

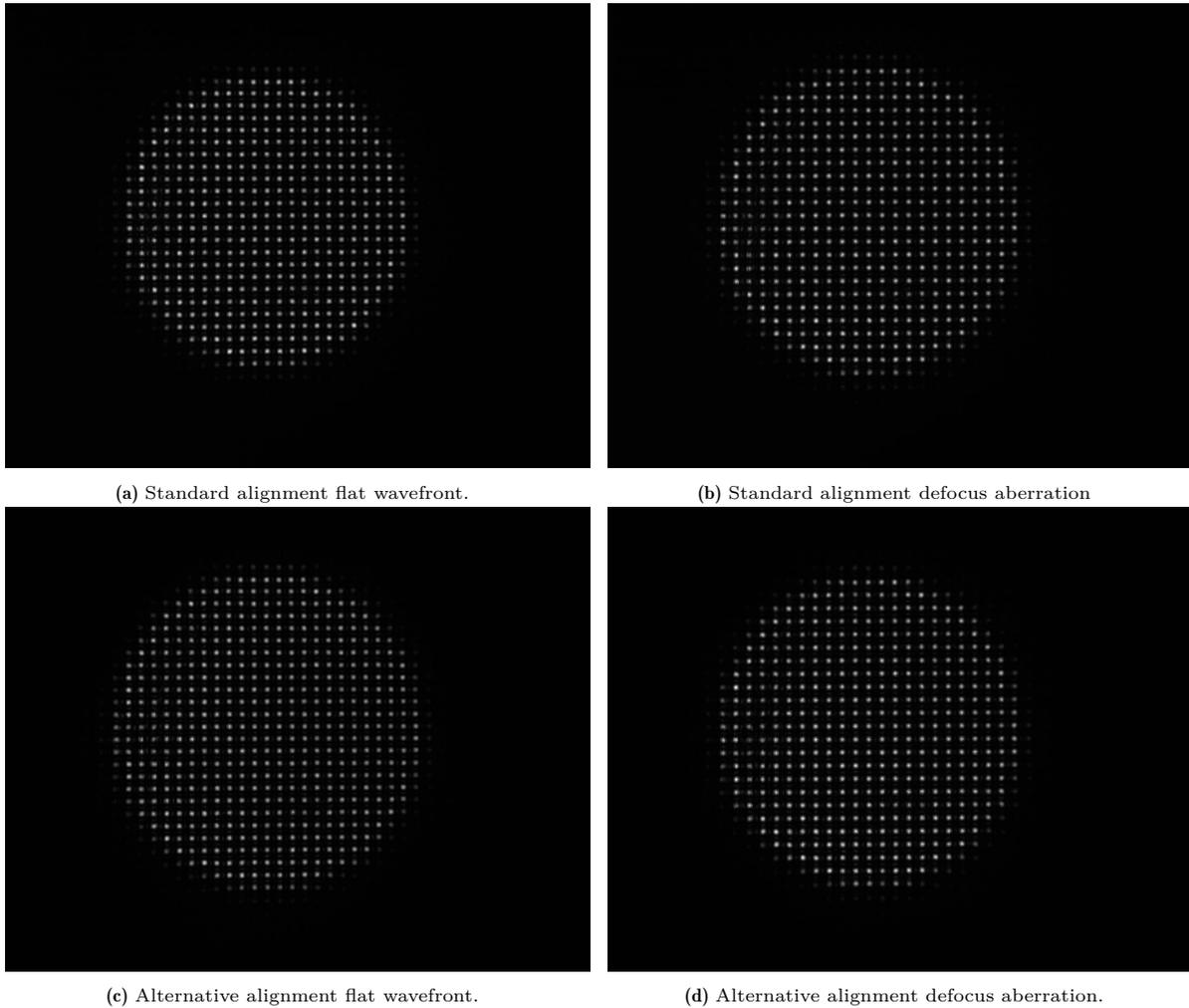


Figure 2.3: Four Shack-Hartmann patterns showing the difference between the alternative and standard alignments. Notice that the total illuminated area on the CCD changes in the case of standard alignment, whereas it remains constant in the case of alternative alignment.

The difference between the resulting Hartmanngrams becomes clear when looking at figure 2.3. The standard alignment patterns will occupy an area on the CCD that changes with the size and shape of the aberration, whereas in the alternatively aligned system the patterns will always occupy the same region, regardless of size or shape of aberration.

2.2. Schematic comparison of alternative alignment with traditional alignment

To illustrate the main difference between standard alignment and traditional alignment, their workings are presented here. First the propagation of the wavefront through the Shack-Hartmann sensor is shown:

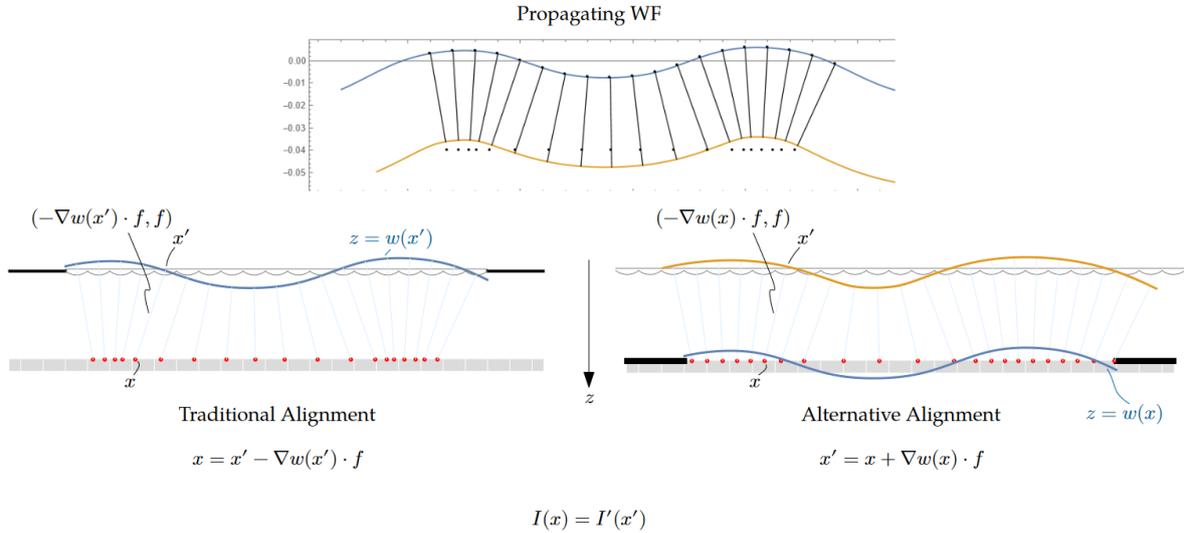
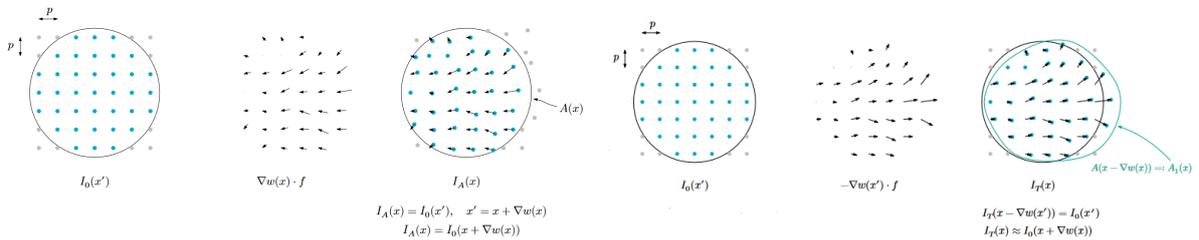


Figure 2.4: Schematic view of the difference between the traditional and alternative alignments. The wavefront propagates from top to bottom. The black bars at x' in the traditional alignment and at x in the alternative alignment denote the locations with optical conjugacy to the deformable mirror.

In figure 2.4 the difference between the traditional alignment and alternative alignment can be seen. An incoming wavefront arrives at the microlens array located at the x' -plane, and is focused on the CCD located in the x -plane. In the traditional alignment the microlens array is optically conjugated to the deformable mirror which is denoted by the black bar. The same black bar is used to denote optical conjugacy between the CCD and the deformable mirror in the alternative alignment.

To calculate the relation between x and x' in the traditional alignment, the light is propagated forward from x . In the alternative alignment the light is propagated backwards from x' to x . The traditional alignment requires the spot shifts to be known at the microlens array, which means the measurement is an approximation. The alternative alignment however requires the aberration to be known at the CCD, which is actually what happens when a Shack-Hartmann wavefront sensor is used. This means that the alternative alignment is not an approximation and should therefore be more accurate. Below is a schematic of the spot patterns for the two alignments.



(a) Schematic view of the spot pattern for the alternative alignment. On the left the reference pattern is shown. In the middle the spot shifts can be seen, and on the right the measured pattern is visible. x' is the location of the lenslet array, and x is the location of the CCD.

(b) Schematic view of the spot pattern for the traditional alignment. On the left the reference pattern is shown. In the middle the spot shifts can be seen, and on the right the measured pattern is visible. x' is the location of the lenslet array, and x is the location of the CCD.

Figure 2.5: Schematic view that shows the comparison of spot patterns for the alternative alignment with the traditional alignment. Notice how the resulting spot pattern for the traditional alignment changes shape depending on the size of the aberration, while retaining the same amount of spots, whereas the alternative alignment does not stray outside the original aperture, but the amount of spots can change.

For the standard alignment the resulting aberrated wavefront can be written as $x = x' - \nabla w(x') \cdot f$, whereas for the alternative alignment it holds that $x' = x + \nabla w(x) \cdot f$, where x is the image plane of the Shack-Hartmann sensor, x' is the location of the lenslets, $w(x)$ and $w(x')$ are the wavefronts at those locations, and f is the focal length of the Shack-Hartmann sensor. There are two important points to be made here. The first is that the region occupied by the spot pattern on the CCD changes its shape depending on the aberration. As the CCD is optically conjugated to the plane of the optical aberration, the region occupied by the spot pattern on the CCD remains constant no matter the aberration. The second has to do with the difference between $\nabla w(x')$ and $\nabla w(x)$. The traditional alignment approximates the actual wavefront:

$$\begin{aligned} I_T(x - \nabla w(x')) &= I_0(x') \\ I_T(x) &\approx I_0(x' + \nabla w(x')), \end{aligned} \tag{2.1}$$

whereas the alternative alignment provides an equality instead of an approximation:

$$\begin{aligned} I_A(x - \nabla w(x)) &= I_0(x') \\ I_A(x) &= I_0(x' + \nabla w(x)). \end{aligned} \tag{2.2}$$

In the case of the standard alignment the lenslet array of the Shack-Hartmann pattern is optically conjugated to the aberration.

2.3. Summary

By placing the MLA in the plane optically conjugated to the aberration alternatively aligned SH sensor is able to measure the curvature of the wavefront without modifying the SH sensor itself. Due to this setup, the region occupied by the spots on the chip no longer changes, but the density of the spots does, whereas in the traditional alignment the region does change. This also means that the alternative alignment is a more accurate way of measuring the wavefront, as it requires the movement of the spots on the CCD, which is provided by the Shack-Hartmann sensor, whereas in the traditional alignment, the movement of spots should actually be measured at the microlens array, which of course does not happen in a Shack-Hartmann sensor. Finally the (scaled) signal from a curvature sensor can also be used directly to actuate bimorph and membrane deformable mirrors as they satisfy the Poisson boundary conditions. This means very little actual control is required to actuate the deformable mirror in a closed loop setting.

3

Using Fourier demodulation to extract the gradients from a Hartmanngram

In the paper titled "Phase retrieval by demodulation of a Hartmann-Shack sensor" [11] Yuval Carmon and Erez Ribak propose to use Fourier demodulation as an alternative to the centroiding method. In their case the displacement of the spots can be seen as a modulation of the pattern as a whole, which is proportional to the phase gradient.

An advantage of using Fourier demodulation compared to using centroiding is that the spots no longer need to stay within their predefined subaperture. As the magnitude of the aberration increases the slopes can become too large and the spots can stray outside of their subapertures. This results in a loss of quality of reconstruction for the centroiding method, whereas this is less of a problem for the Fourier method.

3.1. The method

The method models the intensity function of the spot pattern as follows:

$$I(\mathbf{r}) = V(\mathbf{r})\{2 - \cos[k_x x - F\phi_x] - \cos[k_y y - F\phi_y]\}, \quad (3.1)$$

where $V(\mathbf{r})$ is equal to the pattern amplitude at location $\mathbf{r} = (x, y)$, k_x and k_y are the scaled inverse lenslet pitches: $k_x = 2\pi/P_x$ and $k_y = 2\pi/P_y$. F is equal to the focal length of the lenslet array, and ϕ_x and ϕ_y are equal to the phase gradients in x- and y-direction respectively. This is a simplification from modeling the spots as airy patterns, where instead they are modeled as repeated sinusoids. Using Euler's identity the cosines can be written as complex exponentials, resulting in the following equation:

$$\begin{aligned} I(\mathbf{r}) &= 1/2\{2V(\mathbf{r}) - C_x(\mathbf{r})e^{ik_x x} - C_x^*(\mathbf{r})e^{-ik_x x} - C_y(\mathbf{r})e^{ik_y y} - C_y^*(\mathbf{r})e^{-ik_y y}\}, \\ C_x(\mathbf{r}) &= V(\mathbf{r})e^{-iF\phi_x} \\ C_y(\mathbf{r}) &= V(\mathbf{r})e^{-iF\phi_y}, \end{aligned} \quad (3.2)$$

where F is the focal length of the lenslets in the Shack-Hartmann sensor and the superscript $*$ denotes the complex conjugate. Taking the Fourier transform of this expression results in the following:

$$\begin{aligned} \hat{I} &= 1/2\{4\hat{V} - \hat{C}_x * \delta(q_x - k_x) - \hat{C}_x^* * \delta(q_x + k_x) \\ &\quad - \hat{C}_y * \delta(q_y - k_y) - \hat{C}_y^* * \delta(q_y + k_y)\} \\ &= 1/2\{4\hat{V} - \hat{C}_x(q_x + k_x, q_y) - \hat{C}_x^*(q_x - k_x, q_y) \\ &\quad - \hat{C}_y(q_x, q_y + k_y) - \hat{C}_y^*(q_x, q_y - k_y)\}, \end{aligned} \quad (3.3)$$

where the $\hat{}$ superscript denotes the Fourier transform, and the normal script Asterix denotes a convolution. From equation (3.3) it can be seen that \hat{I} is composed of five parts: \hat{V} , \hat{C}_x and its complex

conjugate \hat{C}_x^* , and \hat{C}_y and its complex conjugate \hat{C}_y^* . As can be seen from equation (3.2), only $C_x(\mathbf{r})$ and $C_y(\mathbf{r})$ and their complex conjugates contain information of the gradients of the wavefront ϕ_x and ϕ_y . As the complex conjugate $C^*(\mathbf{r})$ in general does not contain any information not already contained in $C(\mathbf{r})$ only two terms are required to calculate the wavefront gradient in x- and y-direction: $\hat{C}_x(q_x + k_x, q_y)$ and $\hat{C}_y(q_x, q_y + k_y)$.

To find the wavefront gradient information encapsulated in $\hat{C}_x(q_x + k_x, q_y)$ the lobe is shifted to the Fourier origin by performing a translation of $-q_x\hat{\mathbf{x}}$. A low-pass filter is then applied which preserves $\hat{C}_x(q_x + k_x, q_y)$ and removes all other lobes. Looking at equation (3.3) it can be seen that from this remaining lobe at the center of the Fourier origin the phase information can be extracted by isolating the argument, which results in $F\phi_x$. The same can be done with the y-lobe: shifting $\hat{C}_y(q_x, q_y + k_y)$ by $-q_y\hat{\mathbf{y}}$, applying again the low-pass filter that removes all other lobes, and extracting the phase results in $F\phi_y$.

As the method involves Fourier transforming the lobes padding is used to increase the amount of frequency bins, resulting in a better frequency resolution, which in turn leads to a better reconstruction of the wavefront. To extract the gradients from the Hartmanngram using the method of Ribak and Carmon, one performs the following steps: (1) zero-pad the Hartmanngram to at least twice its size; (2) calculate its Fourier transform; (3a) shift the x-lobe to the origin and (3b) shift the y-lobe to the origin; (4a) apply a low-pass filter that removes all lobes except the central x-lobe and (4b) central y-lobe; (5a) extract the phase of the x-lobe and (5b) y-lobe; (6a) set the padded parts of the resulting x-phase and (6b) y-phase to zero; (7) apply a correction such that the two gradients fields are consistent with one another; (8) extend the corrected gradient fields beyond their aperture; (9) Fourier transform the extended and corrected gradient fields; (10) perform Fourier integration; (11) extract the real part of its inverse Fourier transform and (12) perform phase unwrapping when necessary.

3.2. Summary

To summarise, the method of Ribak and Carmon relies on calculating the Fourier transform of the Hartmanngram and isolating the resulting sidelobes to calculate the gradients. This method is more robust for larger amplitude aberrations than the standard method of centroiding as spots that leave their subaperture no longer automatically pose a problem.

4

Calculating the gradients of a known field using the finite difference

As the Shack-Hartmann sensor gives information on the gradients of the wavefront, this section provides an investigation in the different ways in which a scalar field F can be differentiated. Later it is also shown that the way the derivative is calculated directly influences the method of integration required to go from gradients back to the original scalar field. Given a certain scalar field F different methods of differentiation can be used to arrive at its gradients or slopes. It is shown how the finite difference is calculated in the case of a fully known scalar field F and also a scalar field F that is seen through an aperture.

4.1. Calculating the finite difference from F directly

If the scalar field F is known then the gradients can simply be calculated by finite difference the field. There are two situations: in the first situation there is no aperture and there are no unknown phase jumps. In the second situation the field is seen through an aperture and unknown phase jumps appear at the border. Both of these situations are examined.

4.1.1. Case without aperture

In the case without an aperture the task at hand is fairly simple. An example scalar field F looks like this:

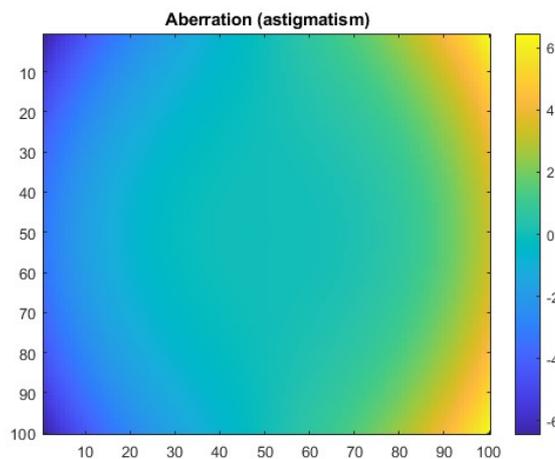
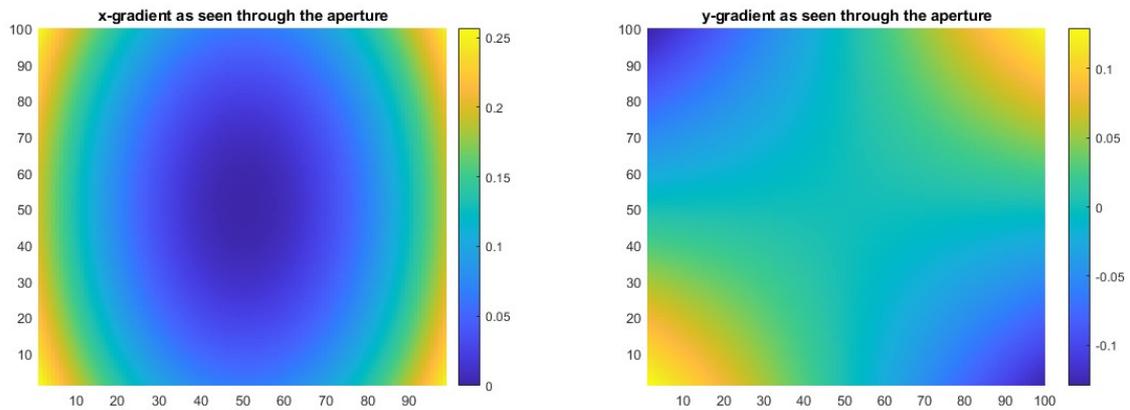


Figure 4.1: Example scalar field F . It is based on astigmatism.

Its gradients can be calculated by taking the finite difference. To take the derivatives the 'dffc' function is used, which calculates the finite difference in the following way:

$$\text{dffc}(F) = [F(2) - F(1) \quad F(3) - F(2) \quad \dots \quad F(n) - F(n-1) \quad F(1) - F(n)] \quad (4.1)$$

Note that the result of $\text{dffc}(F)$ has the same dimensions as the original scalar field F as it also calculates the difference between the first and final pixel in each row or column. The resulting derivatives look like this:



(a) x-gradient of the astigmatism aberration retrieved using the 'dffc' command. (b) y-gradient of the astigmatism aberration retrieved using the 'dffc' command.

Figure 4.2: The x- and y-gradient of the astigmatism aberration. Notice the dimensions of the grids are still 100x100 due to the 'dffc' command.

Integration can now proceed by either summing the rows of one gradient and the column of the other, or by performing Fourier integration. This is covered in the section on integration.

4.1.2. Case with aperture

To illustrate how finite difference would work in the case of an aperture the following scalar field F is given. It is a defocus aberration seen through a circular aperture. In this case there is no information in the white area: its value is not simply equal to 0, but it is set to NaN as there can be no measurements outside of the aperture.

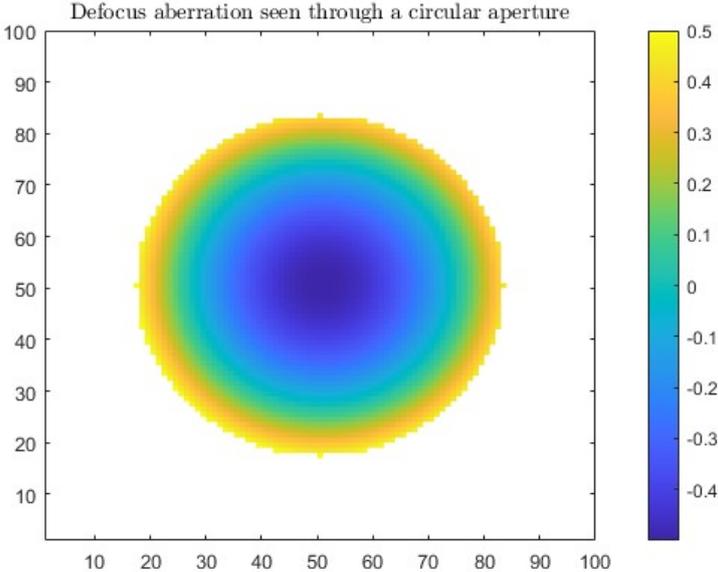
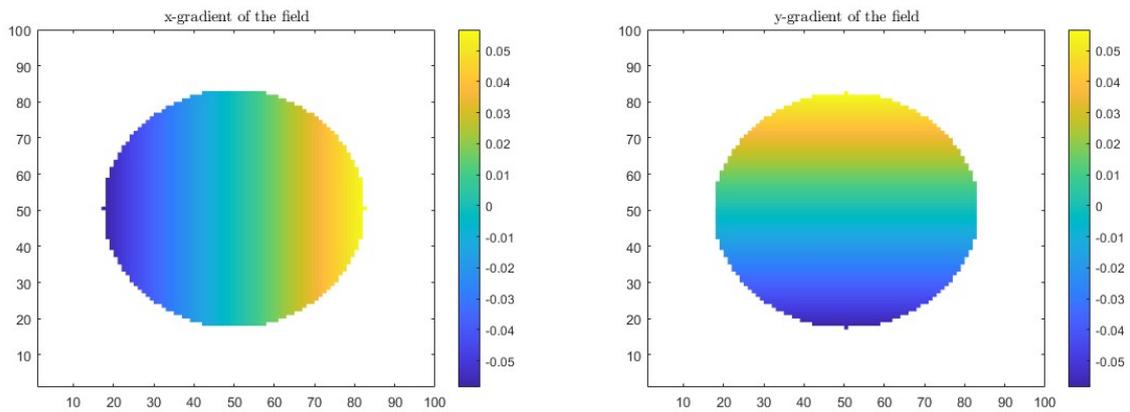


Figure 4.3: Defocus aberration on a 100x100 grid. Its amplitude is equal to +0.5 at the maximum and -0.5 at the minimum. The radius of the aperture is 35 pixels. The values outside the aperture have been set to NaN as there is no information at these points.

The resulting fields in x- and y-direction look like this:



(a) x-gradient of the defocus aberration retrieved using the `difc` command. (b) y-gradient of the defocus aberration retrieved using the `difc` command.

Figure 4.4: The x- and y-gradient of the defocus aberration. Notice the dimensions of the grids are still 100x100 due to the `'difc'` command, but the shape of the circles has been changed slightly due to the NaN values.

These gradients first need to be extended before they can be integrated. This procedure is explained in the section on gradient extension. The subsequent integration is covered in the section on integration.

4.2. Summary

If the scalar field F is known its derivatives can be calculated by using the finite difference method. By using the `'difc'` command the dimensions stay the same. In the case where there is no aperture the gradients themselves are enough to reconstruct the field. In the case of an aperture the two resulting gradients themselves are not enough, due to the unknown jumps at the boundary the field can not be integrated. This is why the gradients need to be extended beyond their original aperture. Chapter 6 shows the process of gradient extension.

5

Calculating and correcting for inconsistencies of a gradient field

In section 3 it was seen that measuring the wavefront using a Shack-Hartmann sensor will result in two components of the gradient field of the phase in perpendicular direction. Because of noise or measurement errors these two components can be inconsistent with one another. In their paper called "Wavefront reconstruction from its gradients" [10], Amos Talmi and Erez Ribak show how to identify and remove the inconsistencies by calculating two correction terms C_x and C_y using an aperiodic Fourier sine series with Dirichlet boundary conditions.

It is important to note that this method supposes the use of finite differencing to calculate the slopes whereas the gradients retrieved using Fourier demodulation do not use finite differencing. Regardless of the method used to calculate the gradient fields, consistency should be enforced anyhow. The different methods of differentiation require different methods of integration, this will be shown in section 7.

5.1. The gradient fields

From the Shack-Hartmann sensor S_x and S_y are measured, which are the gradient fields of the phase Φ . They are defined as follows:

$$\Phi(n+1, m) - \Phi(n, m) \equiv \Delta_x \Phi(n + \frac{1}{2}, m) \approx S_x(n + \frac{1}{2}, m) \equiv S_x(\hat{n}, m) \quad (5.1)$$

$$\Phi(n, m+1) - \Phi(n, m) \equiv \Delta_y \Phi(n, m + \frac{1}{2}) \approx S_y(n, m + \frac{1}{2}) \equiv S_y(n, \hat{m}). \quad (5.2)$$

In these equations Δ is the symmetric difference operator: $\Delta_x A(n + \frac{1}{2}, m) \equiv A(n+1, m) - A(n, m)$. The shorthand notation $\hat{n} = n + \frac{1}{2}$ is used to improve legibility. From the slopes S_x and S_y the wavefront Φ is to be estimated. First

$$\rho(\hat{n}, \hat{m}) = \Delta_y S_x(\hat{n}, \hat{m}) - \Delta_x S_y(\hat{n}, \hat{m}) \quad (5.3)$$

is calculated, which is the mixed derivative. If this is equal to zero then the two gradient fields are consistent with one another, which is equivalent with saying the gradients of the phase are conservative, and $\vec{\nabla} \times \vec{\nabla} \Phi = \vec{0}$. If this is the case then the gradient fields don't require correction and can be used for integration. If this is not the case, the two fields are inconsistent with one another and require correction. A correction is sought which minimises the variance between the actual shears and the measured shears, in other words the minimum of $\epsilon^2 = \sum_{n, m \in R} [\Delta_x \Phi(\hat{n}, m) - S_x(\hat{n}, m)]^2 + [\Delta_y \Phi(n, \hat{m}) - S_y(n, \hat{m})]^2$ should be found.

5.2. Calculation of the correction terms

The correction terms are equal to $C_x(\hat{n}, m)$ and $C_y(n, \hat{m})$ in x- and y-direction respectively. The corrected slopes are equal to

$$\Delta_x \Phi(\hat{n}, m) = S_x(\hat{n}, m) + C_x(\hat{n}, m) \quad (5.4)$$

$$\Delta_y \Phi(n, \hat{m}) = S_y(n, \hat{m}) + C_y(n, \hat{m}), \quad (5.5)$$

meaning that ρ is equal to:

$$\rho(\hat{n}, \hat{m}) = \Delta_y S_x(\hat{n}, \hat{m}) - \Delta_x S_y(\hat{n}, \hat{m}) = \Delta_x C_y(\hat{n}, \hat{m}) - \Delta_y C_x(\hat{n}, \hat{m}), \quad (5.6)$$

as the correction terms are used to counteract the inconsistencies in the gradient fields. The correction terms are sought in such a way that they are derivatives of a certain potential function V :

$$C_x(\hat{n}, m) = -\Delta_y \cdot V(\hat{n}, m) \quad (5.7)$$

$$C_y(n, \hat{m}) = \Delta_x \cdot V(n, \hat{m}).$$

This is done to ensure that the correction terms will be purely solenoidal: V is a pure gradient field and by taking the derivatives this way the resulting correction terms will be pure solenoidal fields. V is now related to ρ in the following manner:

$$\begin{aligned} (\Delta_x^2 + \Delta_y^2) \cdot V(\hat{n}, \hat{m}) &= \Delta_x C_y(\hat{n}, \hat{m}) - \Delta_y C_x(\hat{n}, \hat{m}) = \rho(\hat{n}, \hat{m}) \\ &= \Delta_y S_x(\hat{n}, \hat{m}) - \Delta_x S_y(\hat{n}, \hat{m}) \end{aligned} \quad (5.8)$$

To find V and the corresponding correction terms the geometry of the aperture must be taken into account. Dirichlet conditions are used which means that the potential function should be equal to 0 on the boundaries. For rectangular apertures this means that Fourier sine functions can be used as a potential function. For an $N \times M$ rectangular region the following set of base functions automatically fulfil the boundary conditions:

$$\psi_{k,l}(x, y) = \sin[k\pi(x - \frac{1}{2})/N] \cdot \sin[l\pi(y - \frac{1}{2})/M], \quad (5.9)$$

as it is zero on the boundaries, and a complete orthogonal set over the integer grid $(x, y) = (n, m)$ and the half-integer grid $(x, y) = (\hat{n}, \hat{m})$ excluding the boundary points where the functions are zero.

Expanding the potential function V and the error function ρ in Fourier sine functions yields the following:

$$\begin{aligned} V(\hat{n}, \hat{m}) &= \sum_{q_x=1}^N \sum_{q_y=1}^M \tilde{V}(q_x, q_y) \sin\left(\frac{nq_x\pi}{N}\right) \left(\frac{mq_y\pi}{M}\right) \\ \rho(\hat{n}, \hat{m}) &= \sum_{q_x=1}^N \sum_{q_y=1}^M \tilde{\rho}(q_x, q_y) \sin\left(\frac{nq_x\pi}{N}\right) \left(\frac{mq_y\pi}{M}\right). \end{aligned} \quad (5.10)$$

To transform them back into the standard domain the following equations are used:

$$\begin{aligned} \tilde{V}(q_x, q_y) &= \frac{4}{NM} \sum_{n=1}^N \sum_{m=1}^M V(\hat{n}, \hat{m}) \sin\left(\frac{nq_x\pi}{N}\right) \left(\frac{mq_y\pi}{M}\right) \\ \tilde{\rho}(q_x, q_y) &= \frac{4}{NM} \sum_{n=1}^N \sum_{m=1}^M \rho(\hat{n}, \hat{m}) \sin\left(\frac{nq_x\pi}{N}\right) \left(\frac{mq_y\pi}{M}\right). \end{aligned} \quad (5.11)$$

In Fourier domain they are linked in the following way:

$$\begin{aligned} (\Delta_x^2 + \Delta_y^2) \tilde{V}(q_x, q_y) &= -4 \left(\sin^2\left(\frac{\pi q_x}{2N}\right) + \sin^2\left(\frac{\pi q_y}{2M}\right) \right) \tilde{V}(q_x, q_y) \\ &= -\tilde{\rho}(q_x, q_y), \end{aligned} \quad (5.12)$$

meaning that \tilde{V} can be calculated as follows:

$$\tilde{V}(q_x, q_y) = \frac{\tilde{\rho}(q_x, q_y)}{4\sin^2\left(\frac{\pi q_x}{2N}\right) + 4\sin^2\left(\frac{\pi q_y}{2M}\right)}, \quad q_x, q_y > 0. \quad (5.13)$$

V can now be calculated using equation (5.11), and from equation (5.7) C_x and C_y can be calculated.

5.3. Summary

To summarise, the gradients fields retrieved from measurements using a Shack-Hartmann sensor can be corrupted among other things by measurement errors and noise. This can cause the gradient fields to be inconsistent with one another. These inconsistencies will lead to errors when integrating, therefore a correction is applied to the gradient fields to enforce consistency. The correction terms are purely solenoidal fields aimed at removing all rotation in the gradients. The correction fields are calculated using Fourier sine series using Dirichlet boundary conditions.

6

Extending the gradients beyond their aperture

In the section on the calculation of gradients using the finite difference it was seen that fields seen through an aperture need to be extended beyond that aperture before the gradients can be used for integration. The gradient extension method as shown in "Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform" (2002) by Poyneer, Gavel & Brase is used [6].

6.1. The method of gradient extension

To extend the gradients and resolve the issue of unknown phase jumps there are two conditions that need to be fulfilled: the use of the DFT requires that the resulting wavefront Φ is spatially periodic, which means that the sum of every row in the x-gradient should be equal to zero, and that the sum of every column in the y-gradient should also be zero.

The second condition is that any closed path of the gradient must sum to zero, that is, the gradients need to be conservative. At this point in the process the gradients are already consistent by having added the correction terms C_x and C_y .

All that remains is to extend the gradients by taking the first and final values in a column and extending them all the way up and down respectively for the x-slope, and perform the same operation on the rows and extend to the left and right respectively for the y-slope.

This way of extending automatically guarantees that the sum of any closed path is equal to zero. The only thing that is left is to ensure that all rows in the x-gradient and all columns in the y-gradient sum to zero. This is done by summing all but the last elements of every row in the x-gradient and of every column in the y-gradient, and setting the negative of this result as the value of the last element of every row in the x-gradient and every column in the y-gradient.

Performing gradient extension on an example field looks like this. The conservative gradients of a defocus aberration are shown:

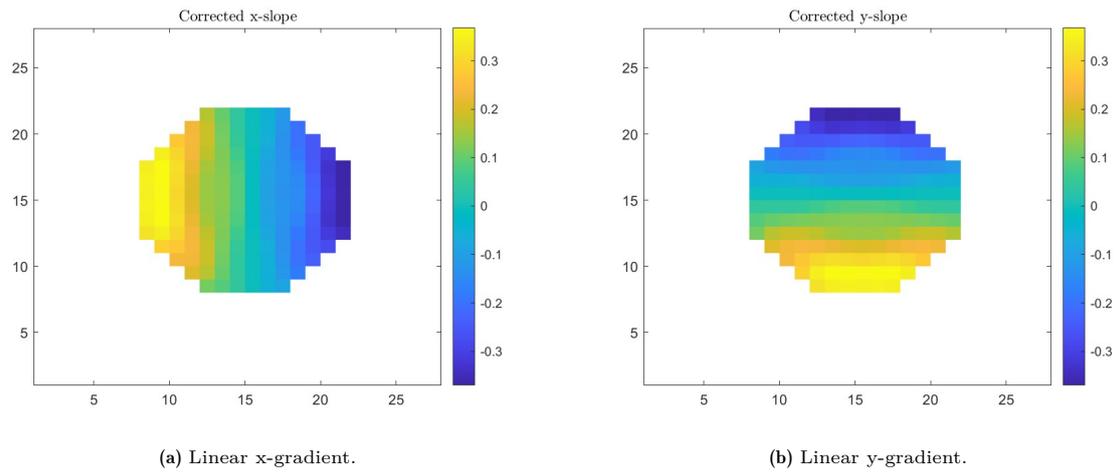


Figure 6.1: Gradients retrieved from a defocus aberration. The white regions are set to NaN values as there is no information from the Shack-Hartmann sensor.

First the gradients are extended and zeroes are put in place of all the remaining NaN values:

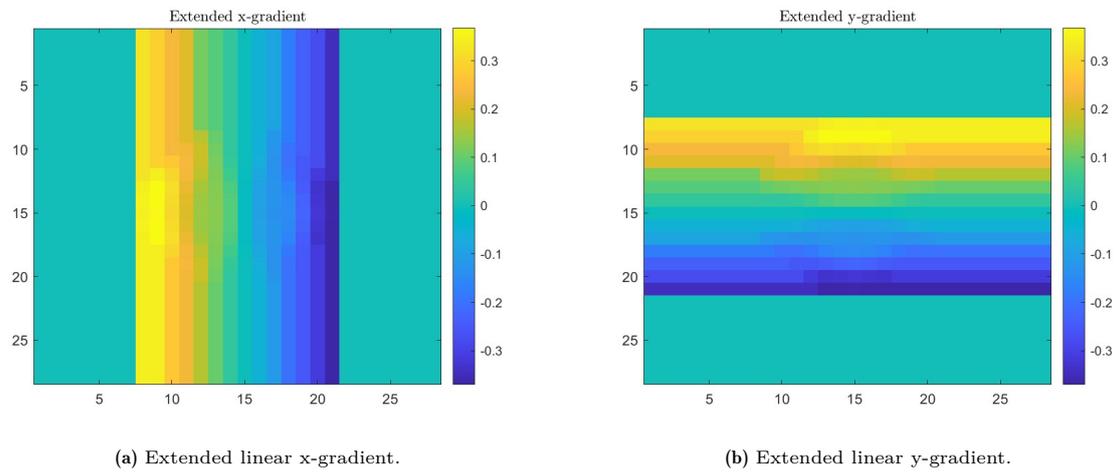


Figure 6.2: Extended gradients retrieved from a defocus aberration. The remaining NaN values after the extension operation have been replaced with zeroes.

After the gradients are extended and the NaN values are replaced with zeroes, the sum of each of the columns and rows can be calculated and the final values can be added:

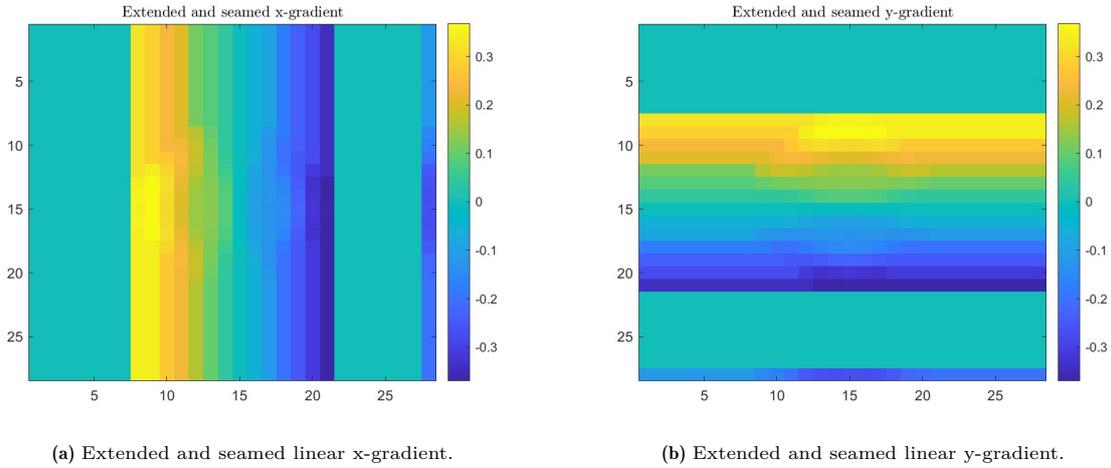


Figure 6.3: Extended and seamed gradients retrieved from a defocus aberration. The rows in x-direction and the columns in y-direction all sum up to zero.

The gradients have now been extended, and the two conditions – 1) that the sum of every row in the x-gradient and every column in the y-gradient should be equal to 0 and 2) that any closed path of the gradient sums to zero – have been fulfilled. The slopes are now ready to be integrated to arrive at the wavefront Φ .

6.2. Summary

It was seen that fields seen through an aperture require extension beyond that aperture in order to be integrated. As a Fourier integration will be used later, the gradient fields need to adhere to two rules: 1) that the sum of every row in the x-gradient and every column in the y-gradient should be equal to 0 and 2) that any closed path of the gradient sums to zero. A simple extension was used: the first and last nonzero value of the rows was copied for the beginning and the end of the row respectively for the x-gradient, and the same is done for the columns in the y-gradient. To ensure the the sum of every row in the x-gradient and every column in the y-gradient equals 0 a seam is added at the right end and bottom end of the x- and y-gradient respectively. This final pixel is equal to the negative of the sum of all the pixels in that row or column which precede it. The gradients are now extended and can be integrated.

7

Integration of the gradients

In chapter 4 it was seen how the derivative of a certain scalar field F can be calculated in different ways. This chapter shows the different integration methods used to arrive at the scalar field F from its slopes. The way the slopes are to be integrated depends on the method that was used to extract the gradients. With the slopes having been corrected and extended it is now possible to integrate and arrive at the wavefront Φ . In this chapter three methods of integration are analysed: numerical integration, integrating using a kernel in the Fourier domain aimed at integrating fields from discrete derivatives, and integrating using a kernel in the Fourier domain aimed at integrating fields from continuous derivatives. It is shown that for an accurate result derivatives calculated using Fourier demodulation require the use of the continuous Fourier kernel.

7.1. The finite difference

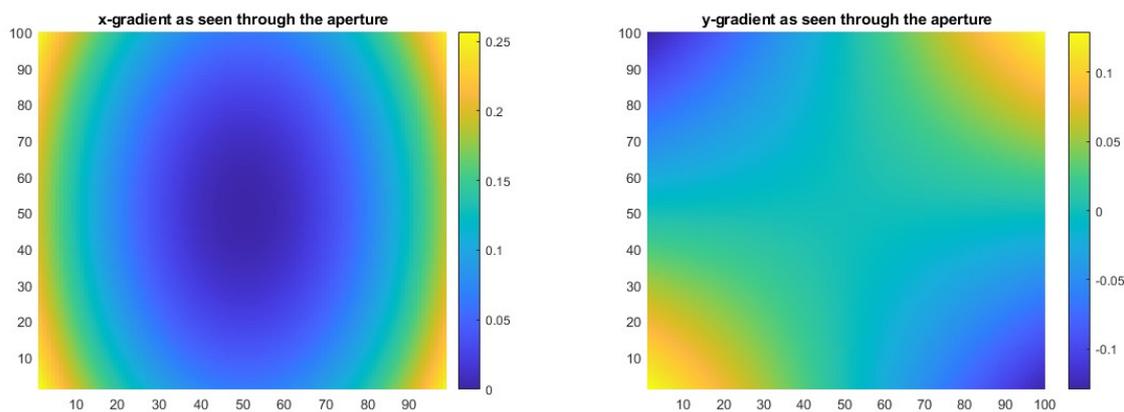
An easy way to calculate the wavefront Φ from the consistent and extended gradients S_x and S_y is to simply sum along the x- and y-gradient fields:

$$\begin{aligned}\Phi(n, m) &= \Phi(1, 1) + \sum_{x=3/2}^{n+1/2} S_x(x, 1) + \sum_{y=3/2}^{m+1/2} S_y(n, y) \\ &= \Phi(1, 1) + \sum_{y=3/2}^{m+1/2} S_y(1, y) + \sum_{x=3/2}^{n+1/2} S_x(x, m),\end{aligned}\tag{7.1}$$

where the initial value $\Phi(1, 1)$ can be any number of choice that results in a piston term. Two methods can be used interchangeably: either the first row is calculated from the x-gradient after which all columns are calculated using the y-gradient, or the first column is calculated using the y-gradient after which all rows are calculated using the x-gradient. As the gradient fields S_x and S_y are consistent the resulting $\Phi(n, m)$ is the same either way.

7.1.1. Case without aperture

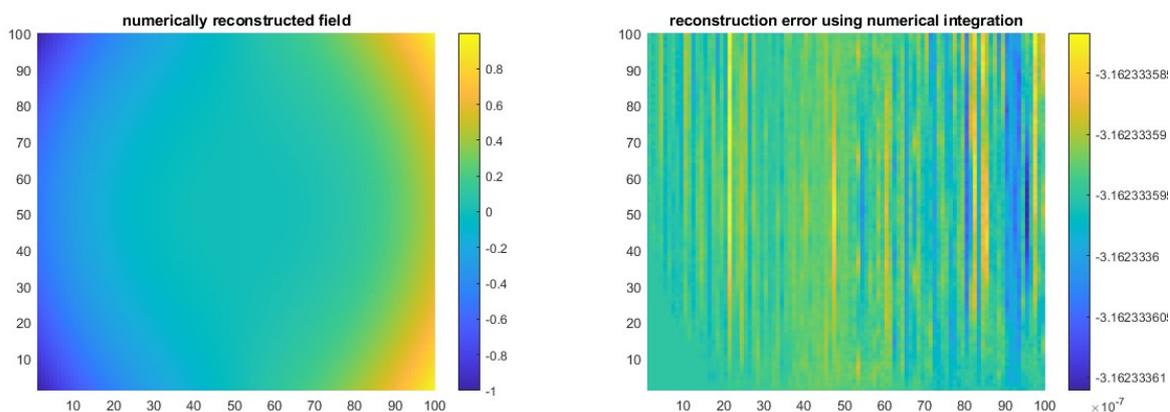
For the case without an aperture the gradient does not need extension. Line integration can be applied directly. Recalling the gradients from the section on the calculation of the slopes:



(a) x-gradient of the astigmatism aberration retrieved using the 'dffc' command. (b) y-gradient of the astigmatism aberration retrieved using the 'dffc' command.

Figure 7.1: The x- and y-gradient of the astigmatism aberration. Notice the dimensions of the grids are still 100x100 due to the 'dffc' command.

Integrating line by line yields the following result:



(a) The reconstructed field from integrating line by line. (b) The difference between the ground truth and the reconstruction.

Figure 7.2: The reconstruction of the field and its reconstruction error. Notice how in the case of a full aperture the reconstruction is near perfect.

As can be seen this method is suitable to integrate fields. In real life the gradients are always measured using an aperture. This simulation of this situation is now considered.

7.1.2. Case with aperture

An example of integrating gradient fields that have been obtained using the finite difference can be seen below:

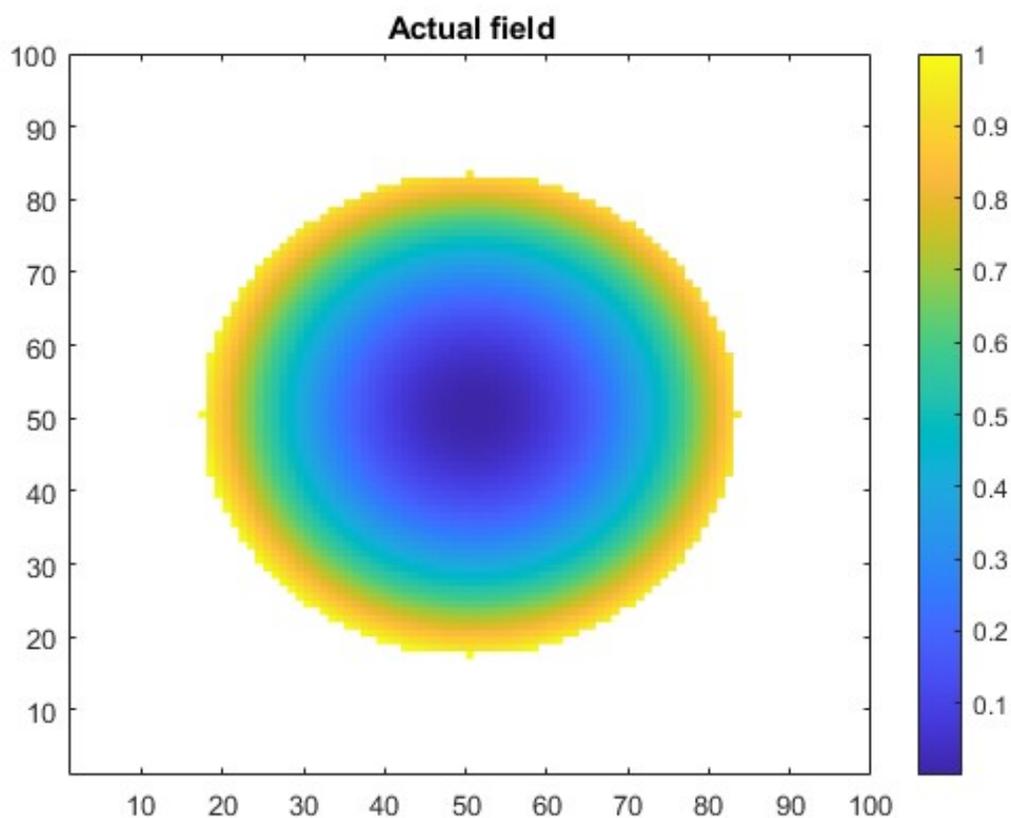
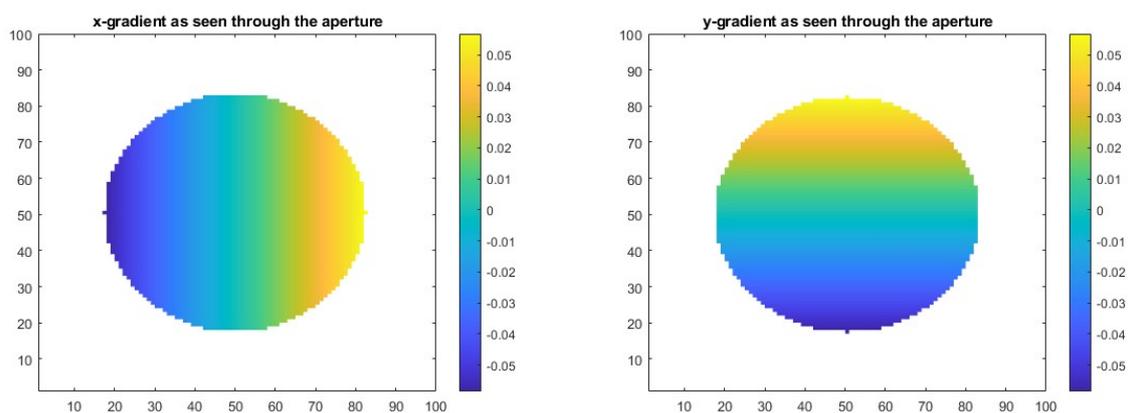


Figure 7.3: Ground truth field for integration.

This field is taken as a basis for the examples. It is a defocus aberration seen through an aperture with a radius of 33 pixels on a 100x100 grid. Taking the derivatives in x- and y-direction yields the following gradient fields S_x and S_y :



(a) Resulting x-gradient from taking the finite difference.

(b) Resulting y-gradient from taking the finite difference.

Figure 7.4: Gradients from the finite difference.

As these gradients have unknown phase jumps at the border of the aperture gradient extension is required. The resulting extended gradients look like this:

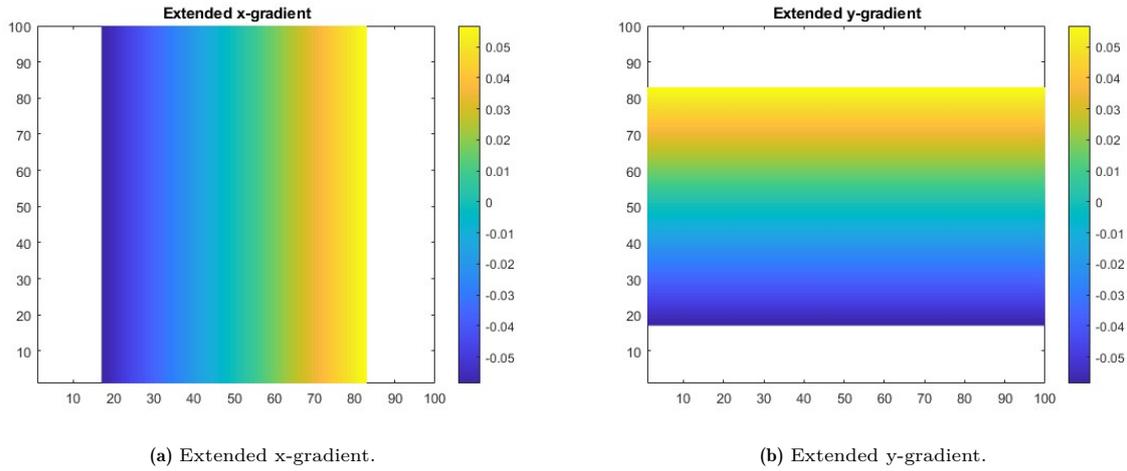


Figure 7.5: The extended x- and y-gradients.

Having extended the gradient, the integration is now possible. For numerical differencing in the case of an aperture the reconstruction and its error look like this:

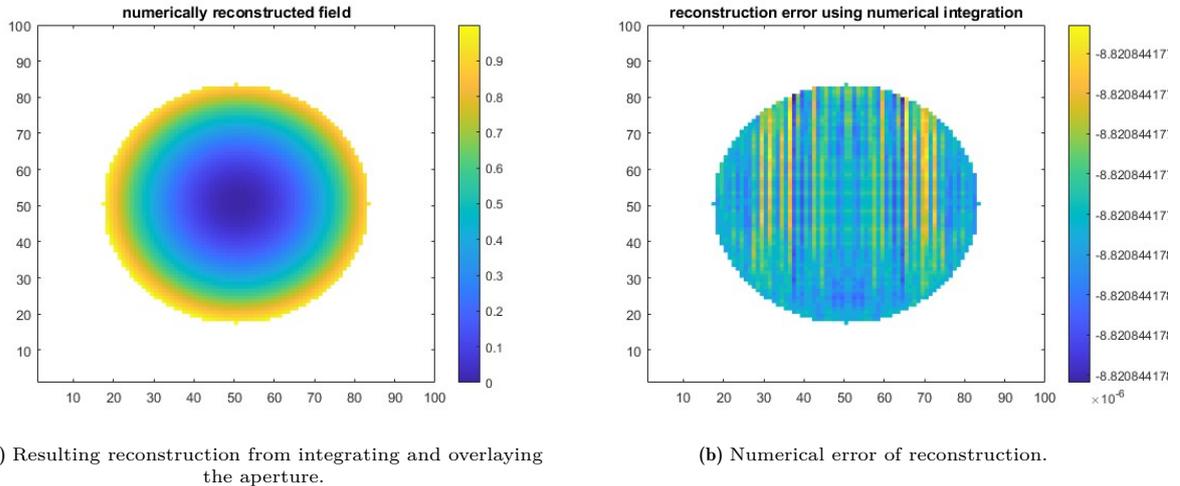


Figure 7.6: The reconstruction and its error using line integration in the case of an aperture.

It can be seen that gradients retrieved from the finite difference can also be integrated using line integration in the case of an aperture.

7.2. Fourier integration

Just as the wavefront Φ can be calculated from its slopes S_x and S_y , the Fourier transform of the wavefront $\mathcal{F}(\Phi)$ can be calculated using the Fourier transforms of the slopes $\mathcal{F}(S_x)$ and $\mathcal{F}(S_y)$. To do this a kernel is to be constructed which relates the three to one another. This kernel depends on the way the derivative is taken: either through the finite difference or through Fourier demodulation. Two kernels are constructed and compared to one another.

7.2.1. Discrete Fourier integration

In the case of the finite difference a discrete Fourier kernel can be constructed to retrieve the Fourier transform of the wavefront from the Fourier transforms of the slopes. To construct the kernel the definition of the finite difference is analysed. The differentiation operation can be seen as follows:

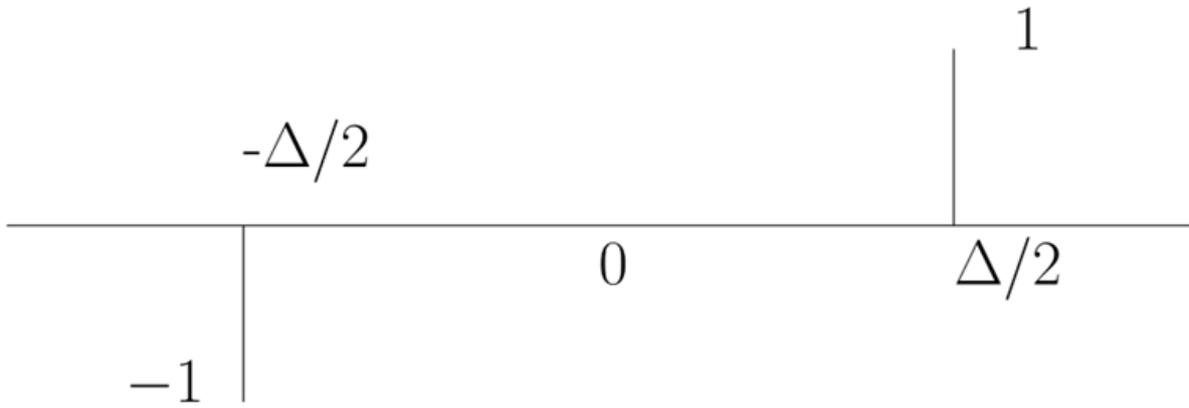


Figure 7.7: A visual representation of the way the derivative $S_x(n + \frac{1}{2}, m) \equiv \Phi(n + 1, m) - \Phi(n, m)$ is taken in one dimension. Δ represents the step size, for real systems Δ represents the distance between the centers of two pixels on the CCD.

The derivative $S_x(n + \frac{1}{2}, m) \equiv \Phi(n + 1, m) - \Phi(n, m)$ is equal to

$$S_x = \left(\delta(t - \frac{\Delta}{2}) - \delta(t + \frac{\Delta}{2}) \right) * \Phi(n, m), \quad (7.2)$$

the convolution of two delta peaks along the phase in x-direction. In the Fourier domain it looks like this:

$$\mathcal{F}(S_x) = \left(e^{-i u \frac{\Delta}{2} 2\pi} - e^{i u \frac{\Delta}{2} 2\pi} \right) \cdot \mathcal{F}(\Phi). \quad (7.3)$$

Using Euler's identity it can be shown that the two complex exponentials are equal to a sine:

$$\mathcal{F}(S_x) = -2i \cdot \sin(u\Delta\pi) \cdot \mathcal{F}(\Phi). \quad (7.4)$$

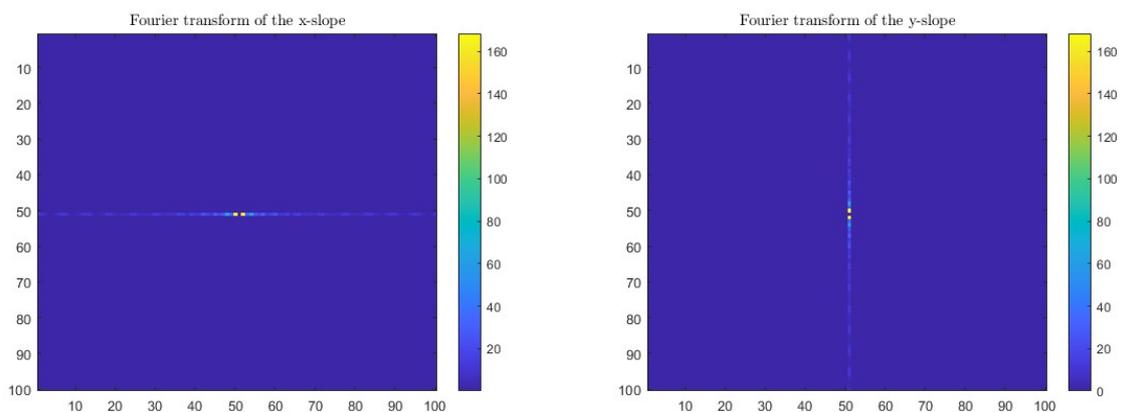
Similarly it can be shown that the gradient field in y-direction is equal to:

$$\mathcal{F}(S_y) = -2i \cdot \sin(v\Delta\pi) \cdot \mathcal{F}(\Phi). \quad (7.5)$$

This means that the following relation holds:

$$\mathcal{F}[\Phi] = i \frac{\sin(u\Delta\pi) \mathcal{F}(S_x) + \sin(v\Delta\pi) \mathcal{F}(S_y)}{2(\sin^2(u\Delta\pi) + \sin^2(v\Delta\pi))}. \quad (7.6)$$

The extended gradients seen in fig. 7.5 can be used directly as the rows in the x-gradient and the columns in the y-gradient sum to 0. Their Fourier transforms look like this:

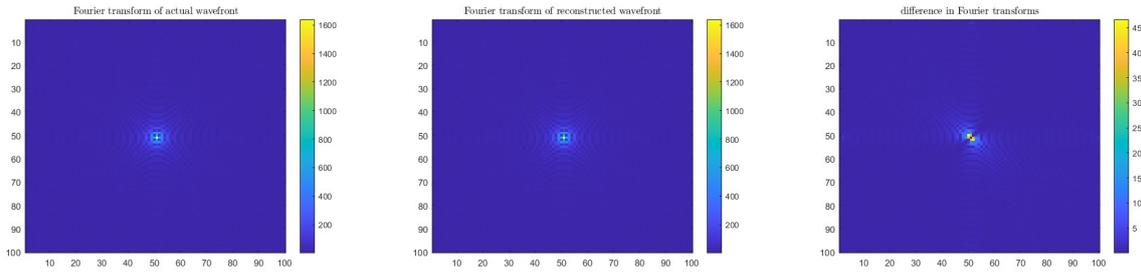


(a) The Fourier transform of the x-gradient.

(b) The Fourier transform of the y-gradient.

Figure 7.8: Fourier transforms of gradients obtained by the finite difference.

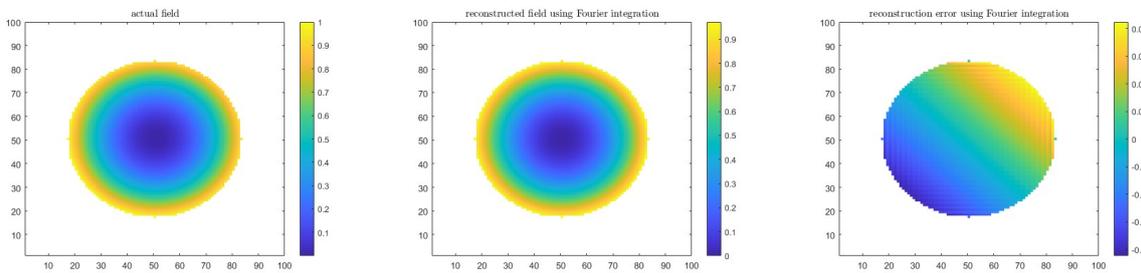
Using equation 7.6 the Fourier transform of the wavefront can be calculated and compared to the Fourier transform of the wavefront:



(a) Fourier transform of the ground truth. (b) Fourier transform calculated using eq. 7.6. (c) The difference between the wavefront and its reconstruction in Fourier domain.

Figure 7.9: The Fourier transform of the ground truth, the reconstruction, and the difference between the two.

By taking the inverse Fourier transform the resulting wavefront can be compared to the ground truth:



(a) The ground truth. (b) Reconstructed field using the discrete Fourier kernel. (c) The difference between the wavefront and its reconstruction in Fourier domain.

Figure 7.10: The ground truth, the reconstruction, and their differences using the discrete Fourier kernel.

As can be seen the discrete Fourier kernel can be used to integrate fields seen through an aperture whose gradient is calculated using finite difference.

7.2.2. Continuous Fourier integration

In the case of a continuous derivative, which is used in the Fourier demodulation method, a different kernel is constructed to retrieve the Fourier transform of the wavefront from the Fourier transforms of the slopes. In the continuous case the derivative is defined as:

$$\frac{d^n f(t)}{dt^n} = (i\omega)^n F(\omega) \quad (7.7)$$

The first order derivative is calculated in the standard domain which is equal to multiplying with $i\omega$ in the Fourier domain:

$$\begin{aligned} \mathcal{F}(S_x) &= i \cdot u \cdot \mathcal{F}(\Phi) \\ \mathcal{F}(S_y) &= i \cdot v \cdot \mathcal{F}(\Phi), \end{aligned} \quad (7.8)$$

where i is the imaginary number and u and v are the x - and y - direction equivalents in Fourier. From this it can be shown that the inverse kernel for the continuous case is equal to:

$$\mathcal{F}\{\Phi\} = \frac{-i \cdot u \cdot \mathcal{F}\{S_x\} - i \cdot v \cdot \mathcal{F}\{S_y\}}{u^2 + v^2}. \quad (7.9)$$

It was seen earlier how the method of Fourier demodulation arrives at gradients from the image of a Shack-Hartmann wavefront sensor. To show how the procedure for the continuous Fourier kernel the starting point are seamed and extended gradients from a defocus aberration:

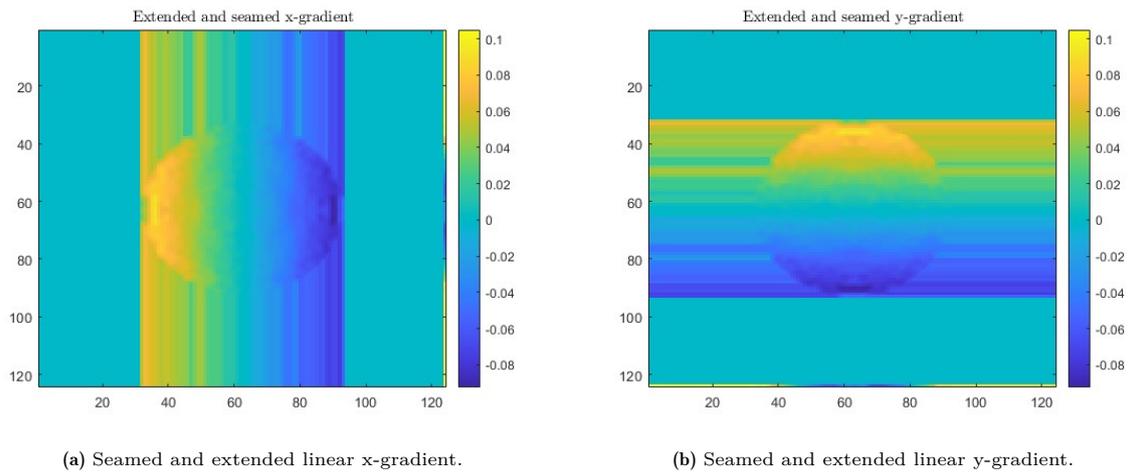


Figure 7.11: Seamed and extended Linear x- and y-gradients retrieved from a Shack-Hartmann wavefront sensor.

Their Fourier transforms look like this:

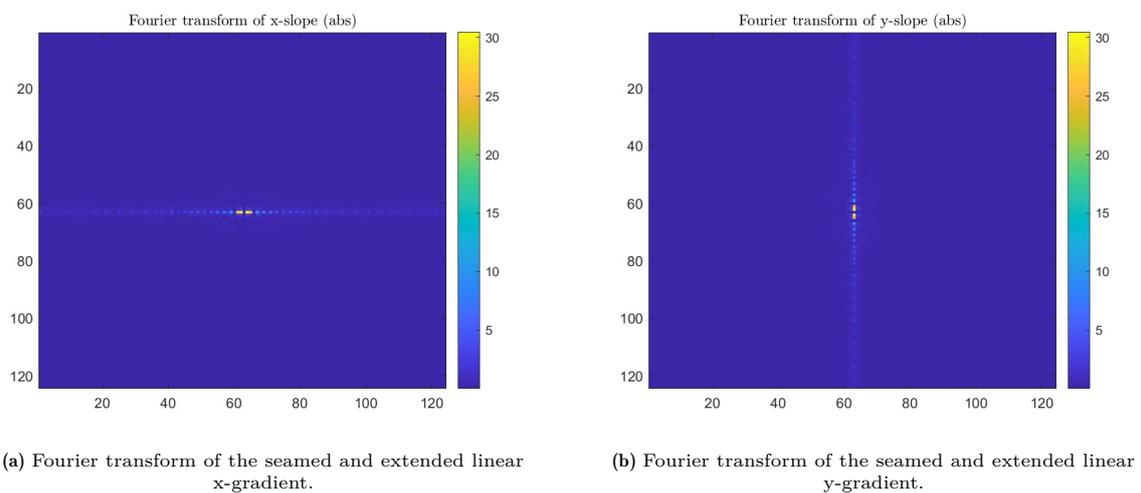


Figure 7.12: Fourier transforms of the seamed and extended Linear x- and y-gradients retrieved from a Shack-Hartmann wavefront sensor.

Using equation 7.9 the Fourier transform of the wavefront can be calculated:

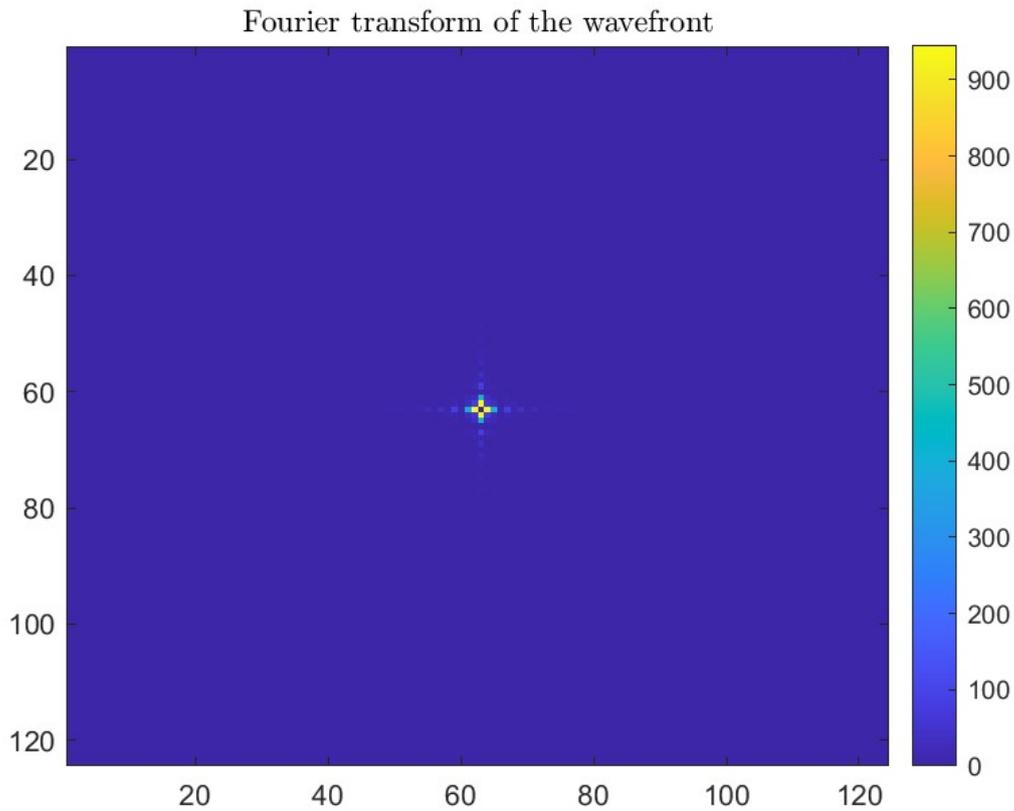


Figure 7.13: Fourier transform of the wavefront calculated using the continuous Fourier kernel.

From this the wavefront itself can be calculated. It is compared to the ground truth:

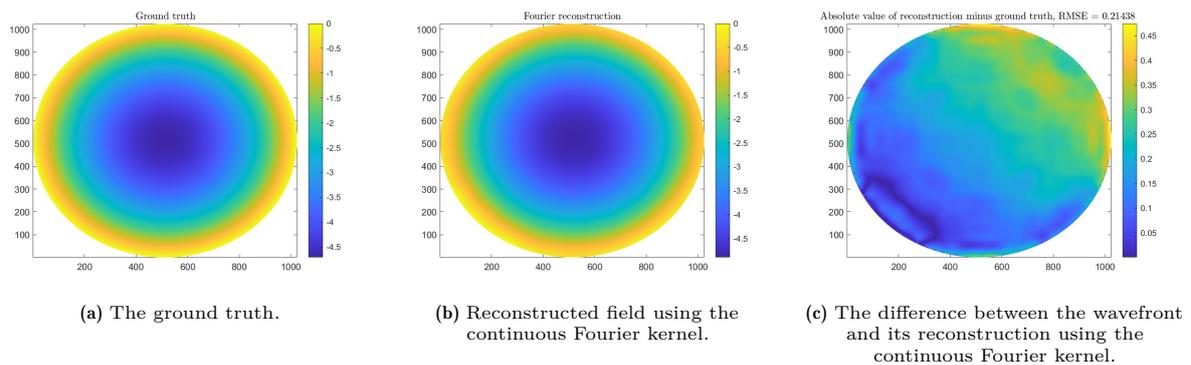


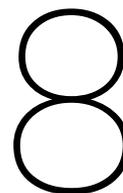
Figure 7.14: The ground truth, the reconstruction, and their differences using the continuous Fourier kernel.

As can be seen the wavefront is reconstructed quite nicely. Only a low RMSE of 0.21 is present.

7.3. Summary

Three methods of integration were shown: line integration and integration using a discrete Fourier kernel in the case of gradients retrieved from finite difference, and integration using a continuous Fourier kernel in the case of gradients retrieved from Fourier demodulation of a Shack-Hartmann pattern. It is shown

that the use of the finite difference in taking the derivatives requires the use of either the line integration or the discrete Fourier kernel, whereas the use of Fourier demodulation in calculating the derivatives requires the use of the continuous Fourier kernel.



Using phase unwrapping to remove phase jumps

Since the combination of the Fourier algorithm and the alternative alignment are able to handle large aberrations it is important that the algorithm is able to deal with wrapped phase. In order to unwrap the phase a slightly modified version of Goldstein, Zebker and Werner's two-dimensional phase unwrapping [9] is used. This algorithm assumes that no phase jump can be equal to or larger than half of a full cycle due to the sampling frequency being at least the Nyquist minimum. The algorithm tries to unwrap the phase field by integrating the phase differences from point to point, always adding an integer number of cycles that minimizes the phase differences.

In the presence of measurement errors and sensor noise an inconsistency can arise where neighbouring points differ by half a cycle or more but adding an integer number of cycles is unable to resolve the situation. If no corrections are made the resulting error propagates globally. This section shows the identification of inconsistencies in the resulting reconstruction, and also what can be done about them.

8.1. The method of phase unwrapping

The situation wherein an inconsistency can not be removed by adding an integer number of cycles shows itself through a circular summation of phase differences. If the sum of the difference over a cluster of four circularly adjacent points is equal to zero the data is consistent. If, however, the data is inconsistent the sum is equal to ± 1 . If it is equal to $+1$ it is called a positive or plus residue, if it is equal to -1 it is called a negative or minus residue. If nothing is done about the residues the resulting reconstruction will suffer from global errors in the form of long streaks in the reconstruction. An example of what happens when a residue is present is given below:

0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4
0.9	0.8	0.6	0.5
0.8	0.8	0.7	0.6

Table 8.1: An inconsistent phase field. All numbers are fractions of a full cycle.

In this field an inconsistency is present in the central four points. This becomes apparent when they are examined more closely:

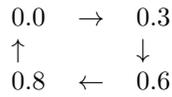


Table 8.2: The inconsistency in the field. As no jumps of more than half a cycle should be possible, the jump from 0.8 to 0.0 should be seen as a +0.2 jump from 0.8 to 1.0 instead of a -0.8 jump from 0.8 to 0.0. Summing clockwise results in a residue of +1.

Upon analysing all clusters it becomes apparent that there is only one residue present in the matrix:

0.0	0.1	0.2	0.3
0	0	0	0
0.0	0.0	0.3	0.4
0	+1	0	0
0.9	0.8	0.6	0.5
0	0	0	0
0.8	0.8	0.7	0.6

Table 8.3: The result of summing all clusters of four points reveals there is only one residue: a positive one in the center. This means there is a net nonzero residue present in the system.

This means that the grid can never be made free of points that are adjacent and differ more than half a cycle. Below is the grid that was seen earlier with the adjacent inconsistencies highlighted:

0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4
0.9	0.8	0.6	0.5
0.8	0.8	0.7	0.6

Table 8.4: An illustration of adjacent points differing more than half a cycle in inconsistent phase field. The two points in red differ more than half a phase, just like the points in orange.

An example of an attempt of removing the inconsistency by adding an integer amount of cycles to the problematic points can be seen below, depending on whether a correction is made from top to bottom or from left to right it results in the following grids:

0.0	0.1	0.2	0.3	0.0	0.1	0.2	0.3
0.0	0.0	0.3	0.4	0.0	0.0	0.3	0.4
-0.1	-0.2	-0.4	-0.5	-0.1	-0.2	0.6	0.5
-0.2	-0.2	-0.3	-0.4	-0.2	-0.2	0.7	0.6

(a) Result of trying to correct the phases by working from top to bottom. (b) Result of trying to correct the phases by working from left to right.

Table 8.5: An illustration of the futility of trying to get rid of inconsistencies in a field by adding or removing one phase. The location of the inconsistency changes (depending even on the direction of correction applied) but the inconsistencies themselves remain.

If an attempt is made to integrate a field where residues are present the local errors will propagate globally, as can be seen below in another more colorful example:

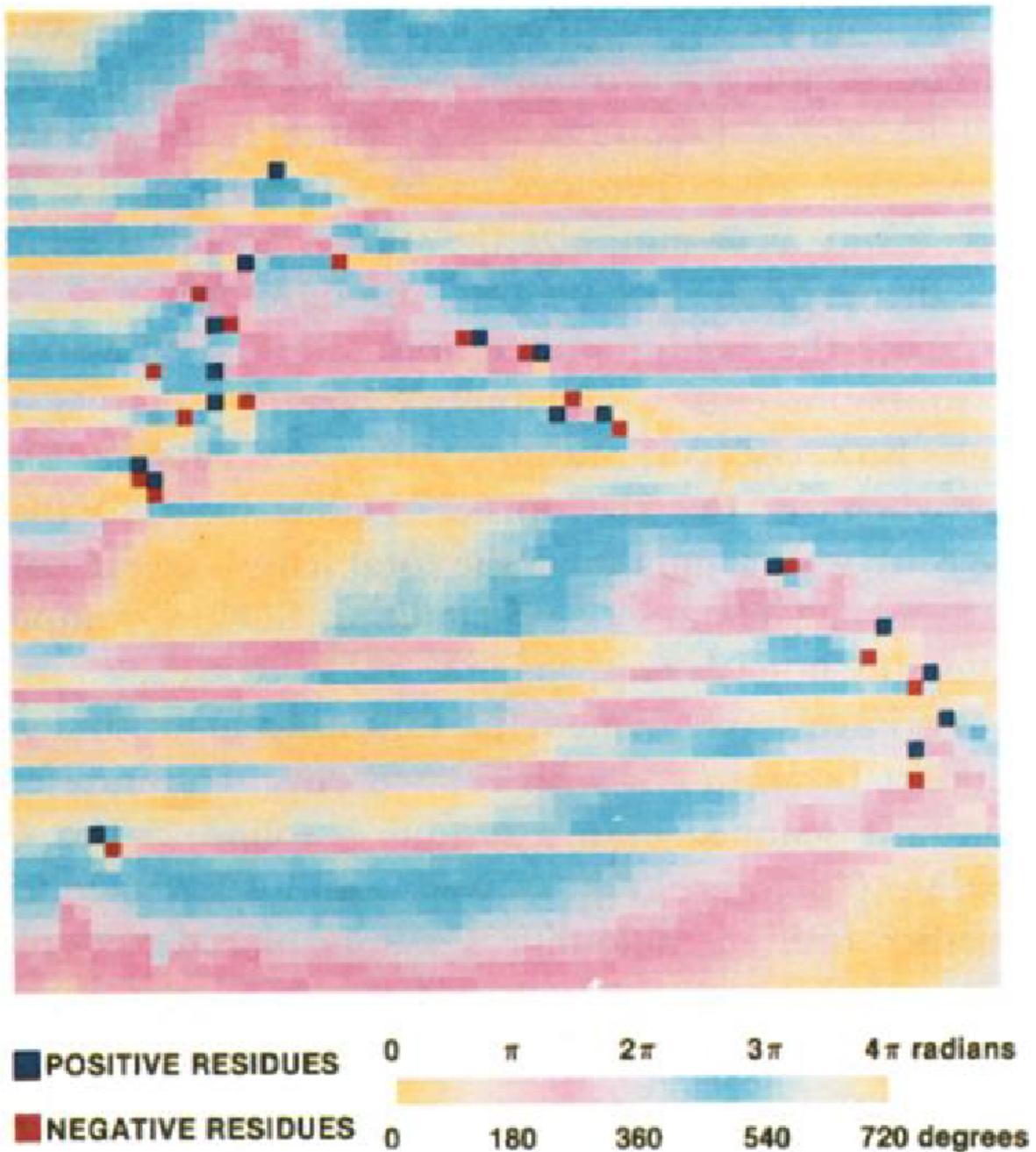


Figure 8.1: An illustration of what happens when the residues are not corrected for. Long lines propagate throughout the solution.

To prevent the residues from propagating globally residues are clustered in such a way that the algorithm does not encircle an area with a net nonzero residue. This is done by mapping the residues and creating branch cuts, which are lines where the algorithm is not allowed to cross. These cuts ensure that no net residues remain by either clustering an equal amount of positive and negative residues together – resulting in a net zero residue – or by isolating the residue through the use of branch cuts, both of which prevent the residues from propagating globally. The downside is that across the lines of the cuts phase discontinuities will arise and local errors will be present. An example of integrating using branch cuts can be seen below:

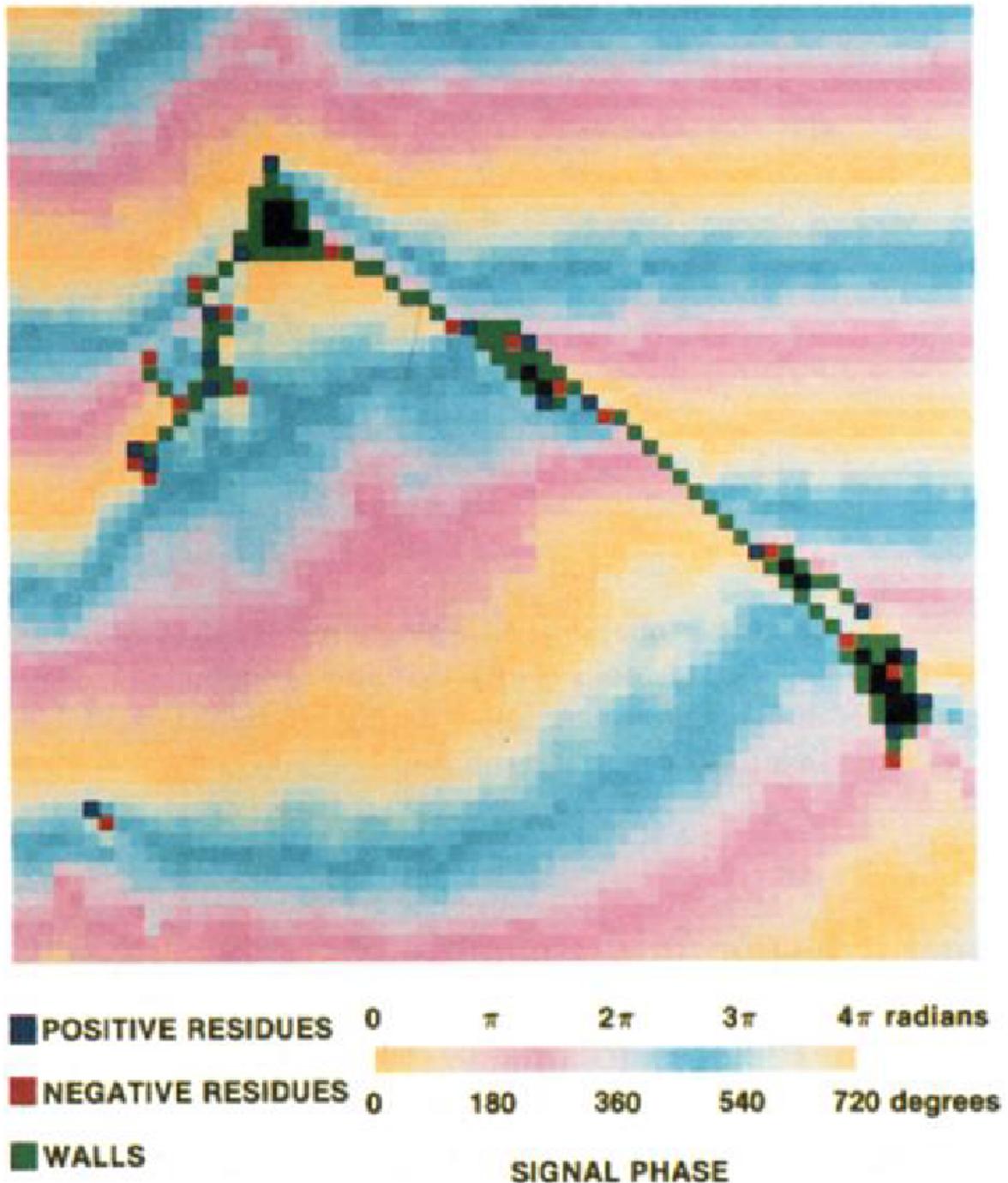


Figure 8.2: An illustration of what happens when branch cuts are put in place to correct for the residues. The residues will not propagate throughout the solution. Compare with figure 8.1 to see the improvement.

The algorithm is a flood fill algorithm. This means that a starting point is defined from which the algorithm sets out. From there it spreads and calculates the phase for all valid adjacent pixels, with the rule that it is not allowed to cross over the branch cuts. If the amount of residues in a certain area is extensive the algorithm might not reach that area as too many cuts have been made. In such cases the algorithm declares the data in that area to be too 'corrupted' to reliably unwrap the phase. This is why the starting location for the algorithm should be located in a large region of the object that can be unwrapped, as starting the algorithm in a noise-rich area might result in it getting trapped. An example of what happens when the phase is too distorted to be unwrapped can be seen below:

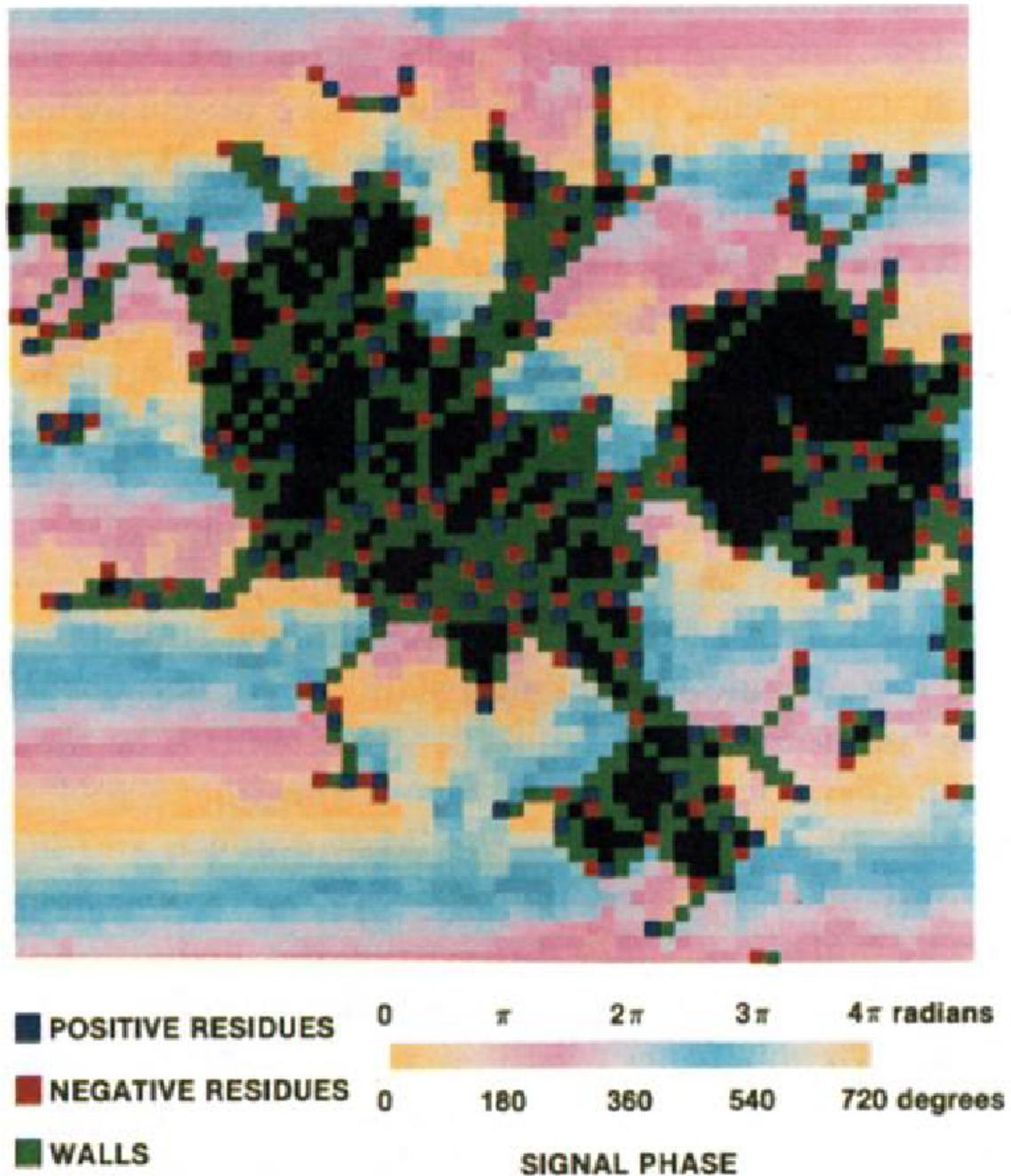


Figure 8.3: An illustration of what happens when the data is noisy to the extent that a large number of branch cuts have to be made: the resulting reconstruction is also mostly corrupted.

8.2. Summary

It is shown that inconsistent fields can be identified by looking for closed loops that sum to either +1 or -1. All the inconsistencies are identified and given the marker of either positive or negative sink. The algorithm then tries to group equal amounts of positive and negative sinks together by building an impenetrable wall around such areas. If the data is corrupted only slightly it is often possible to group the sinks together without the use of lots of these walls. The resulting phase can be unwrapped easily and with high accuracy. In the case of large corruption however it might be that a lot of walls have to be erected to group the sinks or even to block off entire areas. This results in large parts of the resulting phase reconstruction to be illegible.

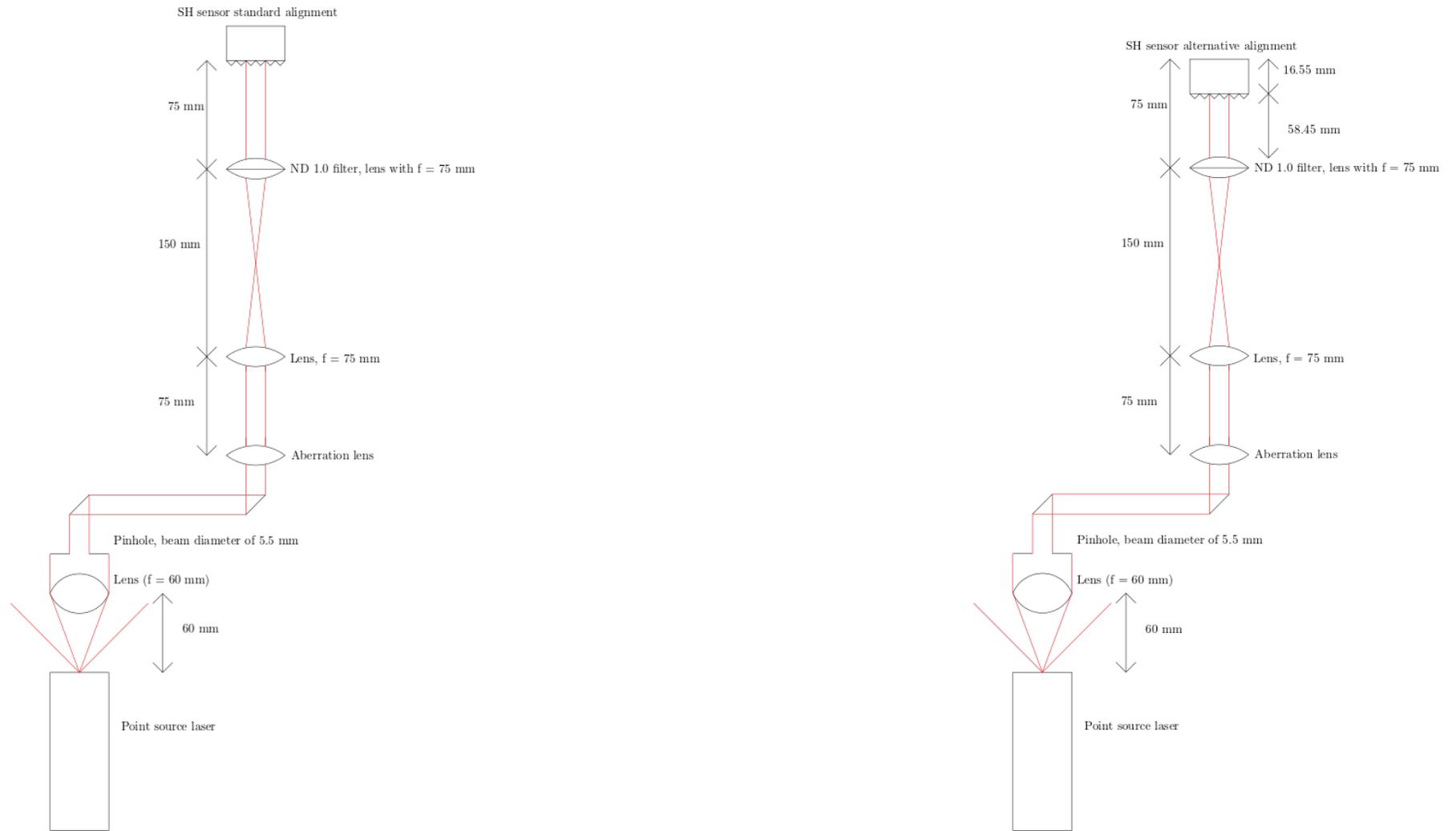
9

Setup for the traditional and alternative alignments

To verify the method works in real life a test setup was created in which known defocus aberrations were created. This section shows a simple setup that was built as a part of the thesis to generate Hartmanngrams.

9.1. Traditional and alternative alignment setups in the lab

The physical experiments consist of introducing some simple known defocus aberrations of different sizes to verify whether the system is aligned correctly and that the algorithm works properly. The setup is depicted below:



(a) Schematic overview of the setup in the standard alignment. The microlens array is placed in the image plane of the system.

(b) Schematic overview of the setup in the alternative alignment. The charge-coupled device is placed in the image plane of the system.

Figure 9.1: Setup of the system. A point source laser with a wavelength of 635 nm emits light that is collimated by a lens. A pinhole regulates the beam diameter of the system. Two mirrors are used to ensure the beam is shot straight at the Shack-Hartmann sensor. Depending on the position of the Shack-Hartmann sensor, the aberration lens is placed either in the plane optically conjugate to the microlens array or the charge coupled device. The Shack-Hartmann sensor is placed on a moving stage in order to switch between the standard and alternative alignment modes.

The Shack-Hartmann sensor is placed on a moving stage which enables an easy switch between the traditional and alternative alignments.

9.2. Hartmanngram generation from the setup

The experiments are conducted as follows: The standard alignment setup is built with the microlens array in the focal plane of the system. A reference image is taken. This provides the flat reference wavefront. After that four lenses, each with a different focal length, are screwed in the aberration lens holder one after the other. These provide four defocus aberrations of different sizes. The lenses have a focal length of 200, 150, 75, and 50 mm. After having captured the four defocus images and the reference, the Shack-Hartmann sensor is moved forward such that the CCD is located at the focal plane of the system, meaning we have shifted the setup into the alternative alignment. The process then repeats: four defocus aberration images and one flat reference image are captured for the alternative alignment as well. These images are then pre-processed, processed, and their results verified.

9.3. Shack-Hartmann wavefront sensor used in the experiments

The Shack-Hartmann wavefront sensor used for the experiments is the model FS3370-O-P300-F18. Its specifications are as follows:

Parameter	Value
Camera model	uEye UI-3370CP-M-GL
Camera type	digital CMOS
Array geometry	orthogonal
Array pitch	300 μm
Array focal length	18.6 mm
Clear aperture	≤ 3.5 mm
Subapertures	≤ 700
Recommended Zernike terms	≤ 300
Wavelength	400 ... 900 nm

9.4. Obtained pictures

The images obtained from the setup can be seen below. Unfortunately they can not be reconstructed as there are three imperfections present in the Hartmanngrams. The first is that the patterns are slightly rotated, which means that the lobes in Fourier space will also be slightly rotated, which makes it impossible to make a nice cutout of the lobe. Second, the illumination is not constant: the intensity changes from spot to spot. The pixel with the highest intensity registers a value close to 256, whereas the center of some other spots barely reaches 60. This intensity change is picked up by the Fourier transform and seeps into the lobe. Third and finally, some residual tip/tilt was still present in the system, which again deteriorates the reconstruction.

In spite of these problems the images are shown nonetheless. The behaviour of the alignments can be seen clearly: in the traditional alignment the spots pattern changes shape but retains the amount of spots, whereas in the alternative alignment the region occupied on the CCD remains roughly the same, whereas the amount of spots decreases with increasing aberration amplitude.

Below are the pictures:

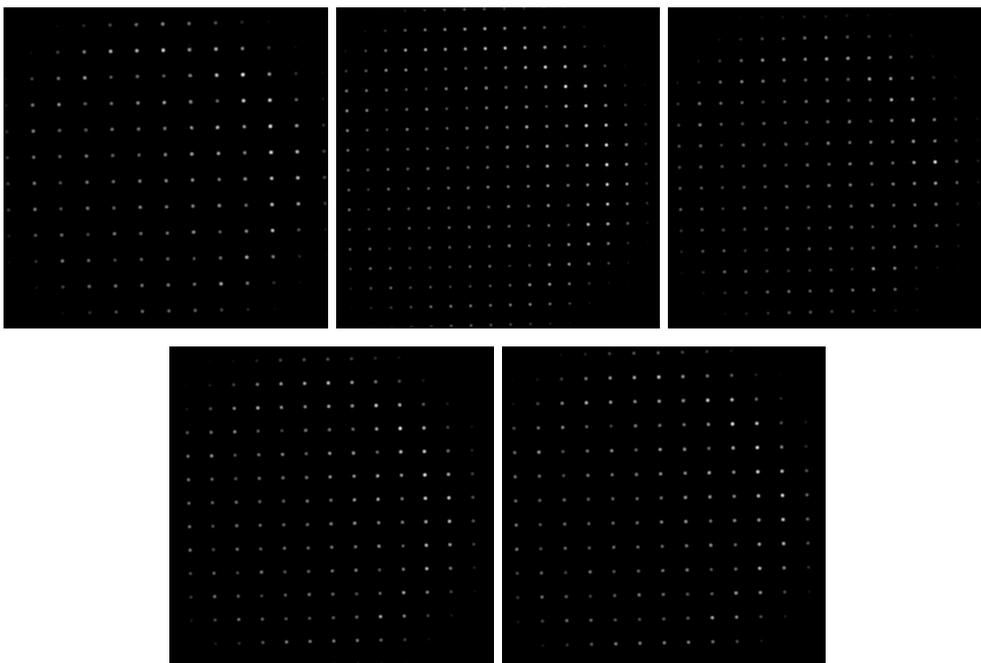


Figure 9.2: Images obtained from the setup in alternative alignment. From top left clockwise the flat reference image, the 50 mm lens, the 75 mm lens, the 150 mm lens and the 200 mm lens.

Above are the pictures obtained by putting traditionally aligning the Shack-Hartmann sensor. Notice how the shape of the pattern changes with the introduction of increasing amplitude aberrations. The amount of spots remains constant throughout all five images.

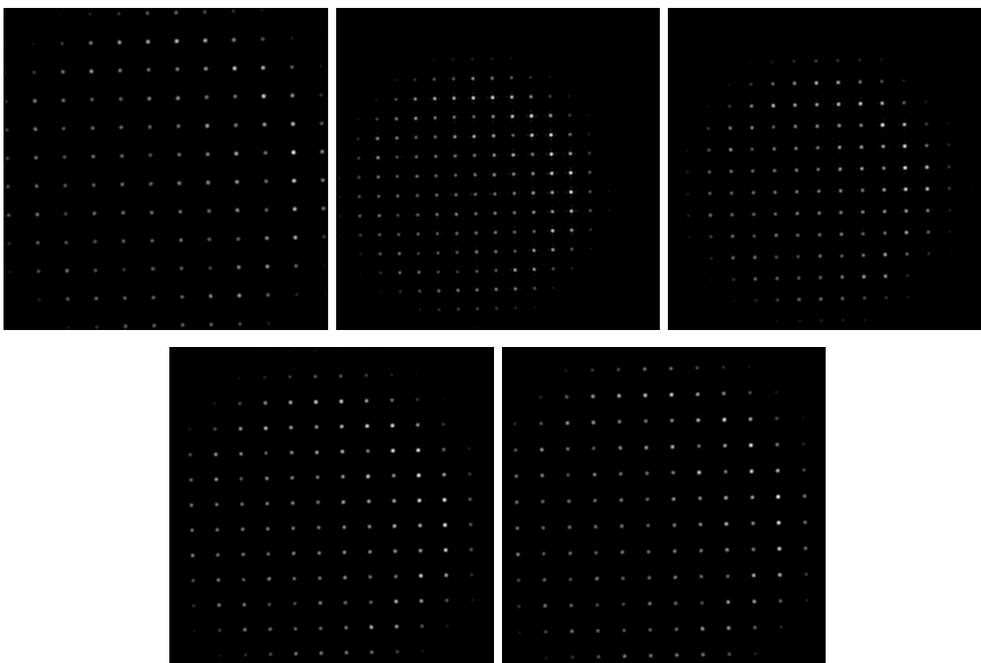


Figure 9.3: Images obtained from the setup in standard alignment. From top left clockwise the flat reference image, the 50 mm lens, the 75 mm lens, the 150 mm lens and the 200 mm lens.

Above are the pictures retrieved from the alternative alignment of the Shack-Hartmann sensor. It can

be seen that the region occupied by the Hartmanngram on the chip remains the same, but now the amount of spots changes with the amplitude of the introduced aberration.

9.5. Summary

A laboratory setup can be achieved with a laser, two mirrors to guide its beam, two lenses to magnify the guided beam, and a Shack-Hartmann sensor to generate the Hartmanngrams. An aberration lens can be placed to introduce various aberrations to the system. The behaviour of the two alignments was demonstrated in reality: The traditional alignment preserves the amount of spots, but the region changes with the aberration, whereas the alternative alignment preserves the region but sees a change in spot count.

10

Results

This chapter shows the results of the combined use of Fourier demodulation with alternative alignment. These results are compared to the centroiding method for both traditional and alternative alignment. The different parameters used for the virtual Shack-Hartmann sensor are discussed. Results are shown for both circular and square apertures, and shot noise (Poisson noise) is also shown. Three types of aberrations are used throughout: positive and negative defocus of the same but opposite amplitude, and Zernike (14,6) to test the algorithms performance on higher order aberrations.

10.1. Simulated Shack-Hartmann parameters

For the simulations Gleb Vdovin's and Fred van Goor's LightPipes for Matlab package was used to generate Shack-Hartmann patterns. As the results are simulated the ground truth is known. The pitch and focal length are varied to see what is the optimal spot pattern for the combined use of the alternative alignment and the Fourier demodulation. In one setup the light is approximated in a more geometrical way. The focal length is short – only 0.5 mm – and the lenslets are close to one another; the pitch is equal to 64 μm . This produces a Hartmanngram that consists of lots of tiny spots. In the other setup the focal length is long, equal to 30 mm, and the lenslet have more space in between: the pitch is equal to 300 μm , This produces a Hartmanngram containing far fewer but significantly larger spots. These spots also interact with one another: the effects of diffraction can be observed, whereas it can not be in the case of the short focal length and small distance between the lenslets. The two situations are shown in table 10.1.

Situation	Focal length (mm)	Microlens array pitch (μm)
Physical optics	30	300
Geometrical optics	0.5	64

Table 10.1: Table showing the focal length and pitch used in the two different scenarios.

10.2. Results using a circular aperture

The results for the circular aperture are shown for four scenarios, each with three aberrations. The four scenarios are as follows:

1. Geometrical optics, alternative alignment
2. Geometrical optics, traditional alignment
3. Physical optics, alternative alignment
4. Physical optics, traditional alignment

The answer to the following two questions is sought: whether the alternative alignment can outperform the traditional alignment, and whether the geometrical optics setup outperforms the physical optics setup or vice verse. The three aberrations are as follows:

1. $\frac{3\pi}{2}$ amplitude defocus
2. $-\frac{3\pi}{2}$ amplitude defocus
3. $\frac{3\pi}{2}$ amplitude Zernike polynomial (14,6)

It is important to check whether the alternative alignment significantly outperforms the traditional alignment in one of these scenarios, or whether the measured difference remains constant throughout.

10.2.1. Geometrical optics, alternative alignment

In this subsection the results for the alternative alignment using the geometrical Shack-Hartmann parameters are shown. The use of the geometrical parameters in the Shack-Hartmann sensor leads to a grid containing a high density of small spots without the effects of diffraction. These results can be compared against the ones in section 10.2.2 to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.4 to see the difference between the geometrical and physical Shack-Hartmann sensor.

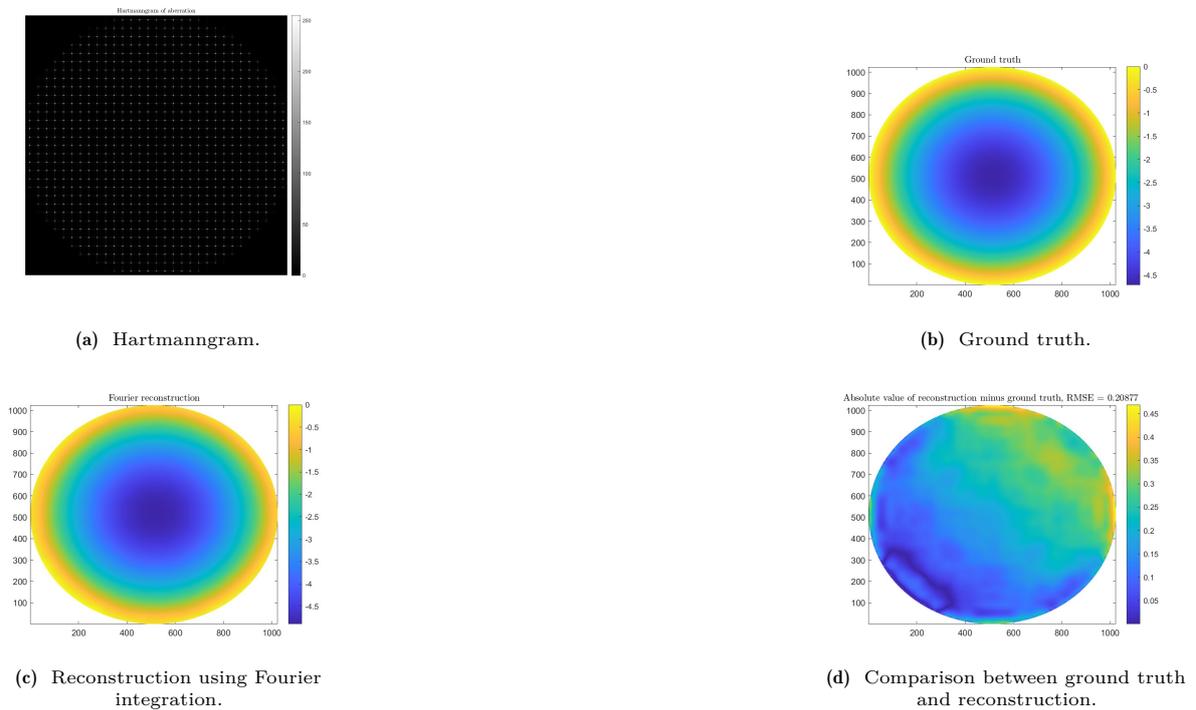
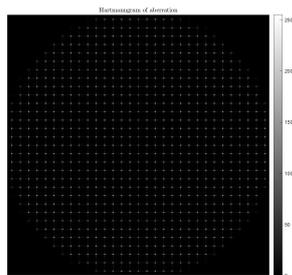
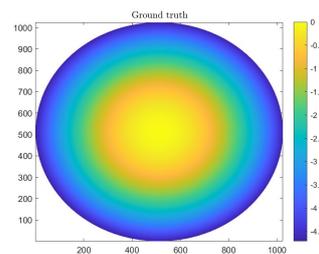


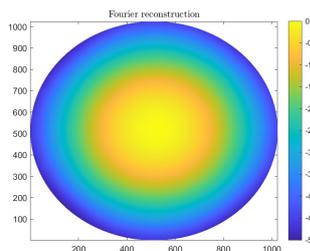
Figure 10.1: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.



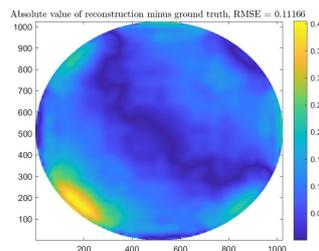
(a) Hartmanngram.



(b) Ground truth.

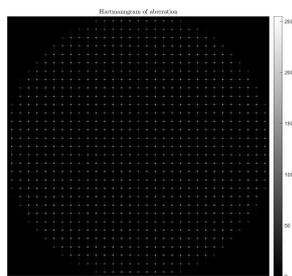


(c) Reconstruction using Fourier integration.

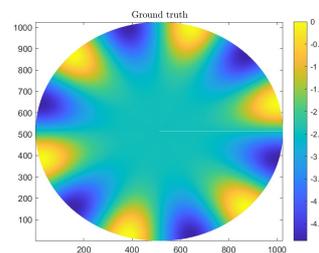


(d) Comparison between ground truth and reconstruction.

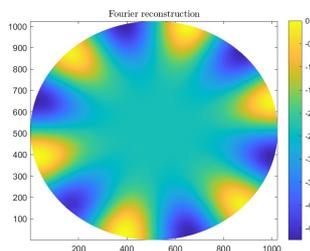
Figure 10.2: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{-3\pi}{2}$ is analysed.



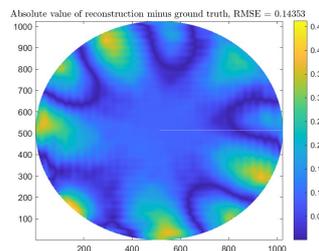
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.

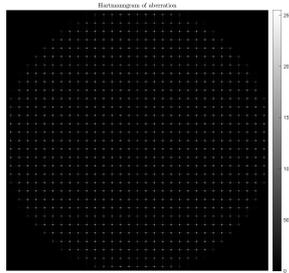


(d) Comparison between ground truth and reconstruction.

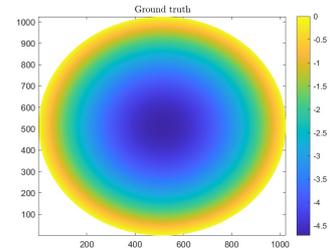
Figure 10.3: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a geometrical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

10.2.2. Geometrical optics, traditional alignment

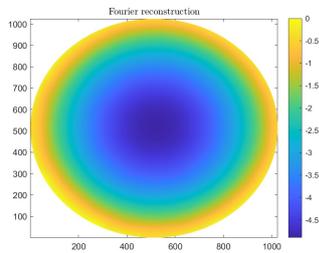
In this subsection the results for the traditional alignment using the geometrical Shack-Hartmann parameters are shown. The use of the geometrical parameters in the Shack-Hartmann sensor leads to a grid containing a high density of small spots without the effects of diffraction. These results can be compared against the ones in section 10.2.1 to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.3 to see the difference between the geometrical and physical Shack-Hartmann sensor.



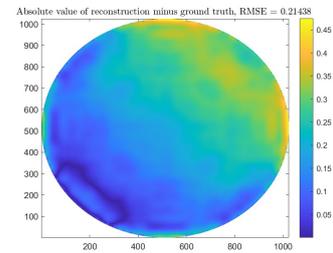
(a) Hartmanngram.



(b) Ground truth.

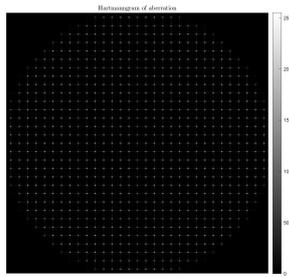


(c) Reconstruction using Fourier integration.

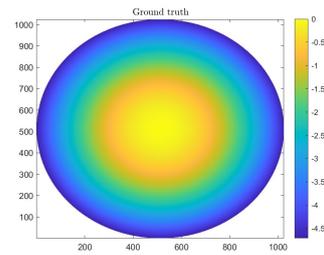


(d) Comparison between ground truth and reconstruction.

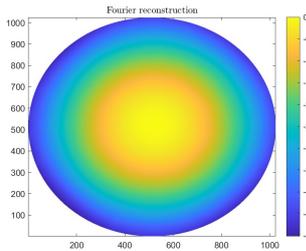
Figure 10.4: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.



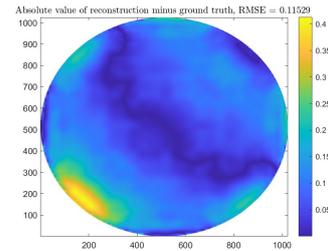
(a) Hartmanngram.



(b) Ground truth.

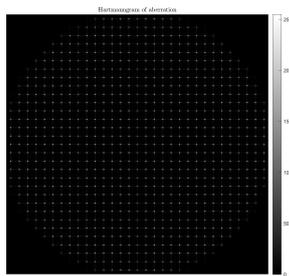


(c) Reconstruction using Fourier integration.

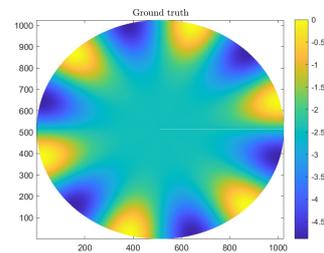


(d) Comparison between ground truth and reconstruction.

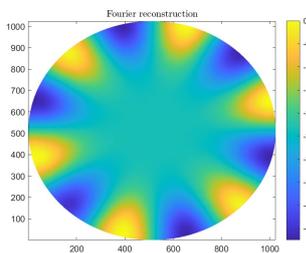
Figure 10.5: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{-3\pi}{2}$ is analysed.



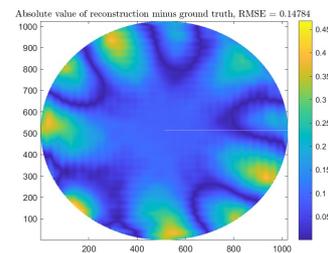
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.

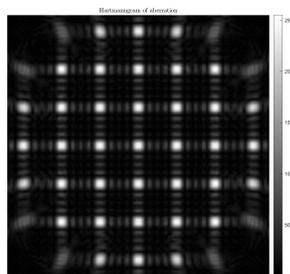


(d) Comparison between ground truth and reconstruction.

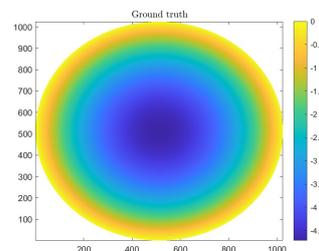
Figure 10.6: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a geometrical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

10.2.3. Physical optics, traditional alignment

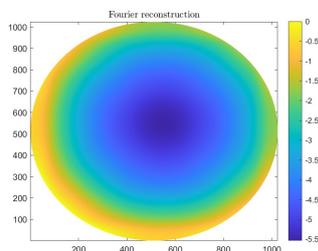
In this subsection the results for the traditional alignment using the physical Shack-Hartmann parameters are shown. The use of the physical parameters in the Shack-Hartmann sensor leads to a grid containing a low density of large spots with the effects of diffraction. These results can be compared against the ones in section 10.2.4 to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.2 to see the difference between the geometrical and physical Shack-Hartmann sensor.



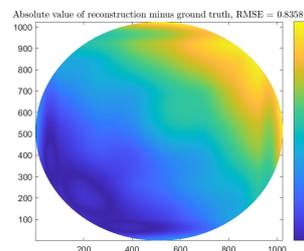
(a) Hartmanngram.



(b) Ground truth.

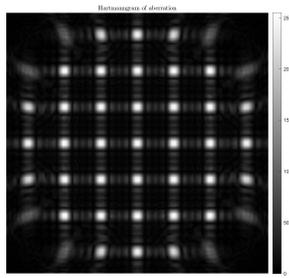


(c) Reconstruction using Fourier integration.

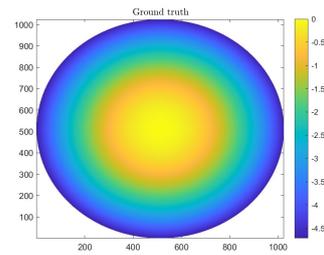


(d) Comparison between ground truth and reconstruction.

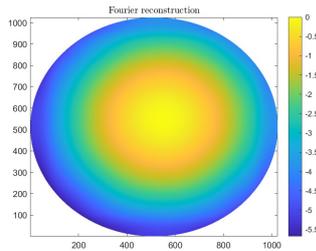
Figure 10.7: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a physical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.



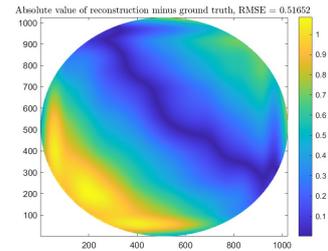
(a) Hartmanngram.



(b) Ground truth.

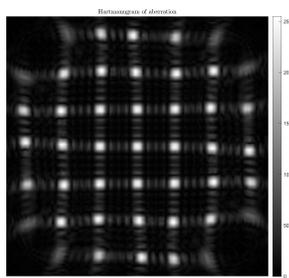


(c) Reconstruction using Fourier integration.

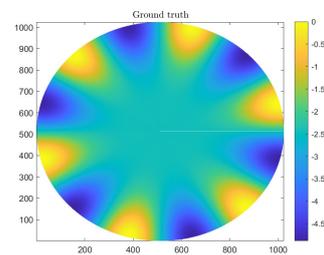


(d) Comparison between ground truth and reconstruction.

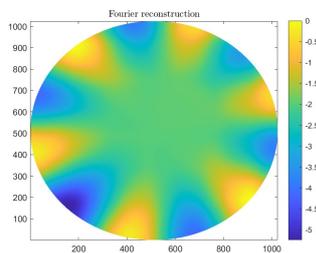
Figure 10.8: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a physical lenslet array. A defocus aberration of amplitude $\frac{-3\pi}{2}$ is analysed.



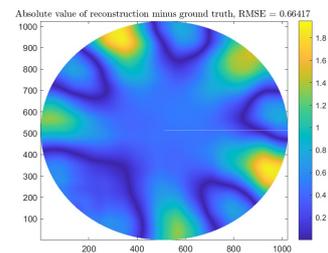
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.

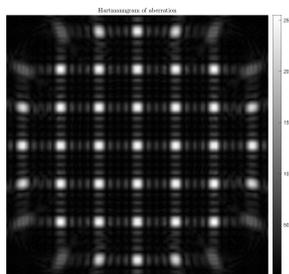


(d) Comparison between ground truth and reconstruction.

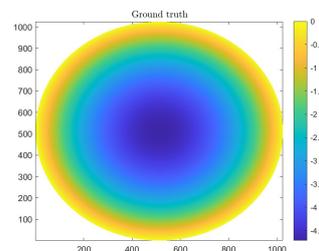
Figure 10.9: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a circular aperture and a physical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

10.2.4. Physical optics, alternative alignment

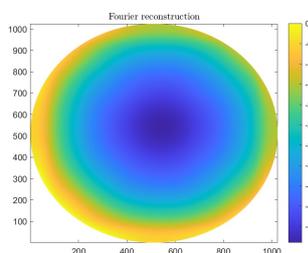
In this subsection the results for the alternative alignment using the physical Shack-Hartmann parameters are shown. The use of the physical parameters in the Shack-Hartmann sensor leads to a grid containing a low density of large spots with the effects of diffraction. These results can be compared against the ones in section ?? to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.1 to see the difference between the geometrical and physical Shack-Hartmann sensor.



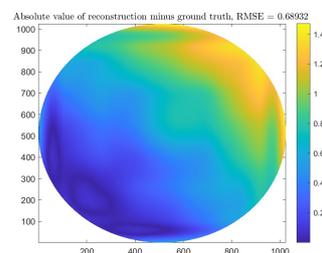
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.



(d) Comparison between ground truth and reconstruction.

Figure 10.10: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a physical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.

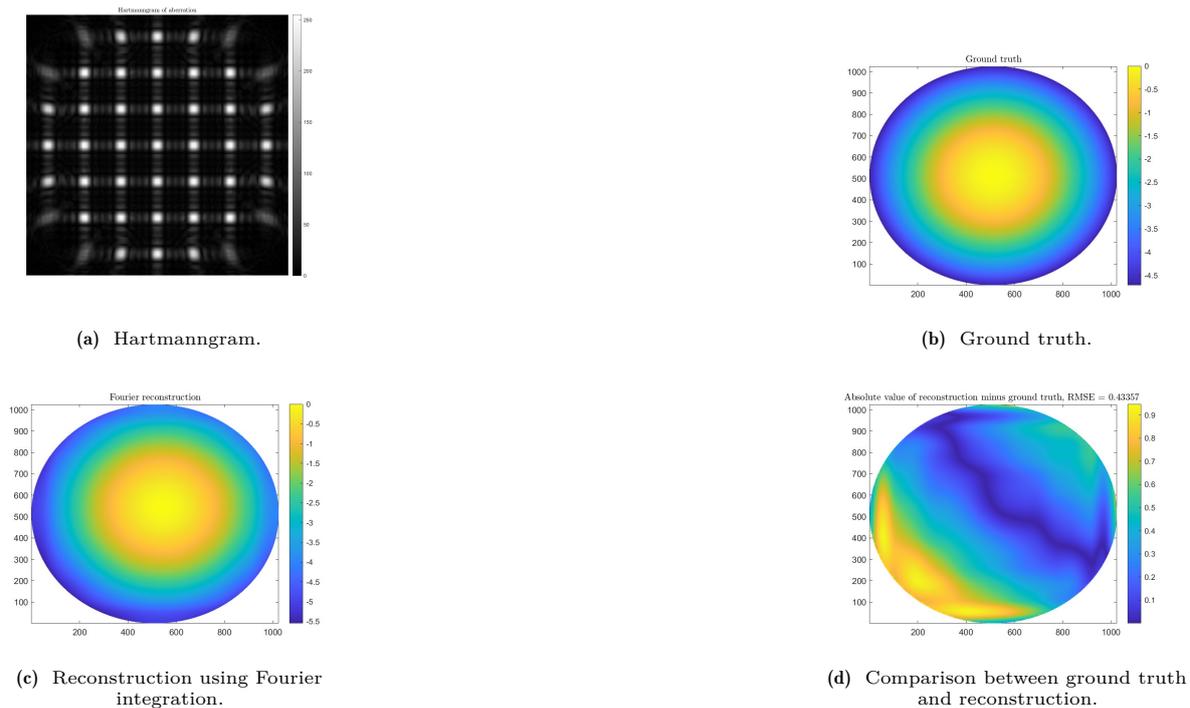


Figure 10.11: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a physical lenslet array. A defocus aberration of amplitude $-\frac{3\pi}{2}$ is analysed.

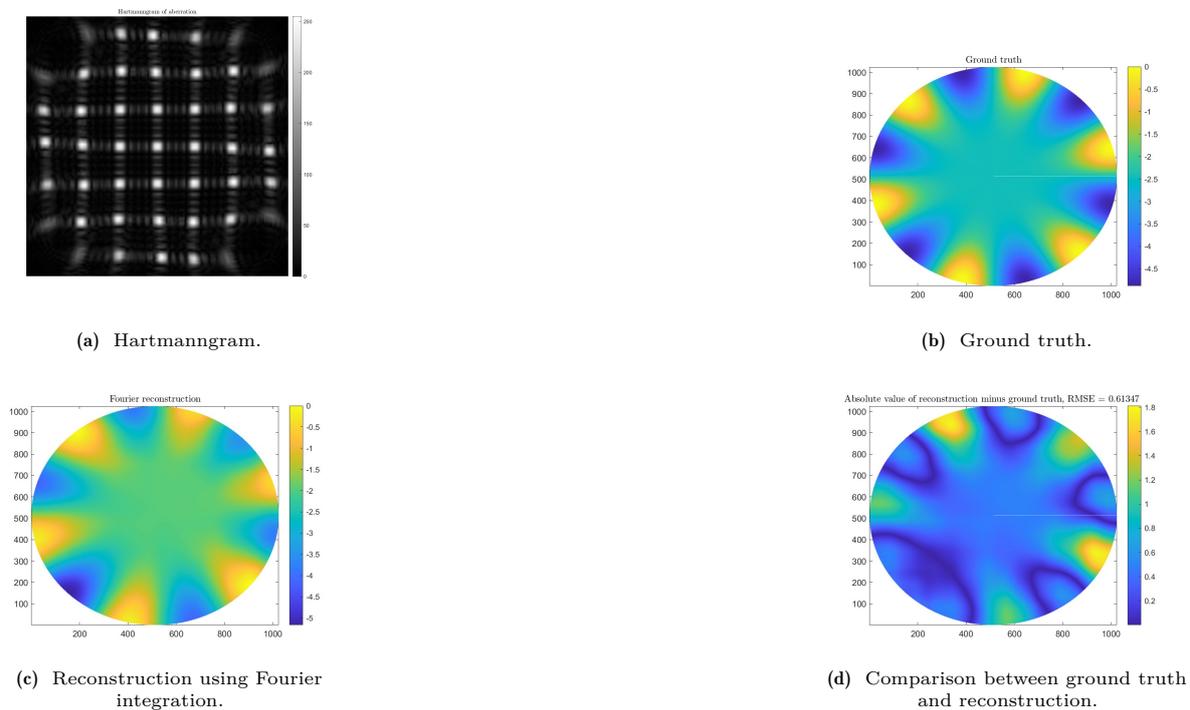


Figure 10.12: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a circular aperture and a physical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

The results can be summarised by a table showing their RMSE values:

	RMSE value
Geometrical optics, alternative alignment	0.20877
positive defocus	0.11166
Negative defocus	0.14353
Zernike (14,6)	
Geometrical optics, traditional alignment	0.21438
positive defocus	0.11329
Negative defocus	0.14784
Zernike (14,6)	
Physical optics, traditional alignment	0.8358
positive defocus	0.51652
Negative defocus	0.66417
Zernike (14,6)	
Physical optics, alternative alignment	0.68932
positive defocus	0.33357
Negative defocus	0.61347
Zernike (14,6)	

Table 10.2: RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.

10.3. Comparison with centroiding

The scenarios that could be compared to FrontSurfer, a commercially available centroiding algorithm, are shown in the table below. The residue is the sum of the absolute values of all coefficients that are not the ground truth coefficient.

	Ground truth coefficient	Ground truth residue	Fourier reconstruction coefficient	Fourier reconstruction residue	FrontSurfer coefficient	FrontSurfer residue
Circular 3π defocus geometrical traditional	2.3364	0.0225	2.3521	0.3604	2.5467	0.7607
Circular -3π defocus geometrical traditional	-2.3364	0.0225	-2.3849	0.3978	-2.5269	0.8644

Table 10.3: Table showcasing the difference in reconstruction coefficients for the Fourier demodulation method versus the centroiding method. Notice how the Fourier reconstruction coefficient is closer to the ground truth coefficient than the FrontSurfer coefficient, and how the residue is small for both circumstances.

10.4. Shot noise

The Shack-Hartmann sensor can suffer from shot noise. This can slightly alter the Hartmanngram and cause the resulting gradients to be noisy as well. It was already seen that two inconsistent gradient fields can be made consistent, this would remove most of the noise. Below is an example where the inconsistencies are not corrected for. This is done to show the robustness of the Fourier demodulation.

Traditional alignment

The example here is the Shack-Hartmann sensor in the traditional alignment with a defocus aberration of 3π . To achieve the noise, the 'imnoise' command is used in Matlab with the 'Poisson' option.

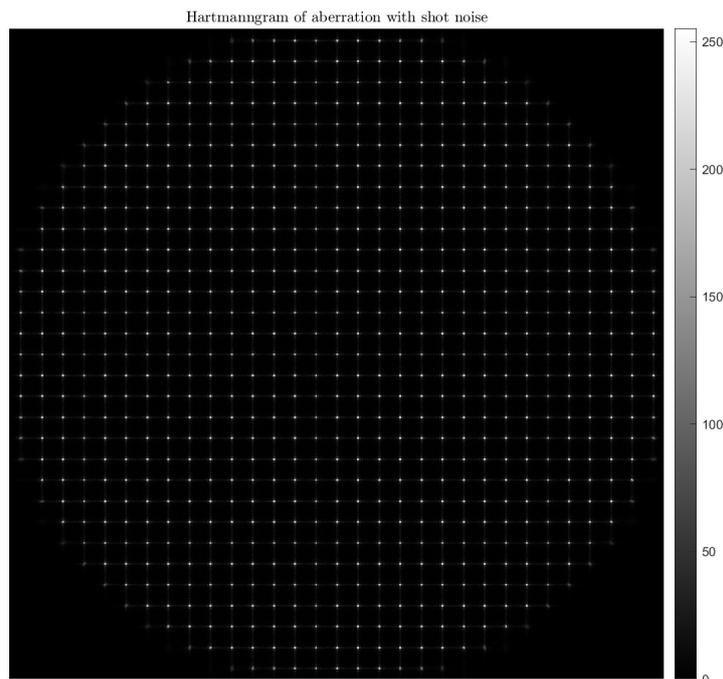
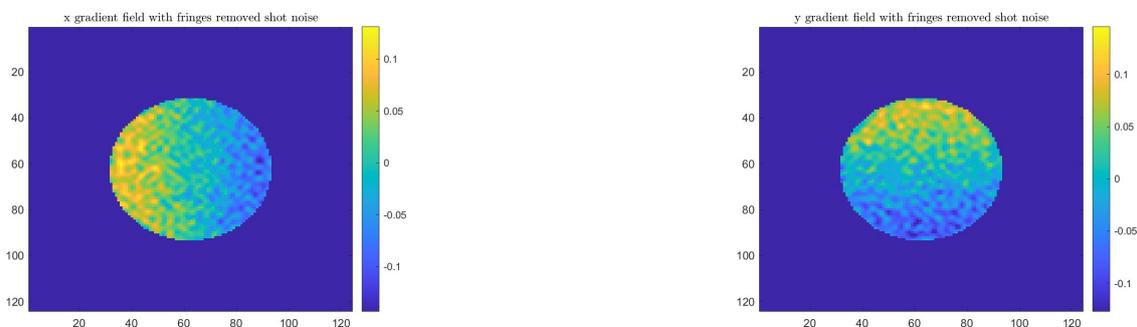


Figure 10.13: Defocus aberration Hartmanngram with shot noise.

Running the Fourier demodulation algorithm on it results in the following gradients:



(a) The resulting x-gradient in the case of shot noise.

(b) The resulting y-gradient in the case of shot noise.

Figure 10.14: The resulting gradients in the case of shot noise on the Hartmanngram. Notice the noise is quite severe, these are gradients of a defocus aberration, but they are no longer nicely linear.

To show the robustness of the algorithm to noise, no correction is made for inconsistent gradient fields. This greatly increases the speed of the algorithm as it is no longer required to calculate the Fourier sine transforms. The resulting wavefront, the ground truth and their difference can be seen below:

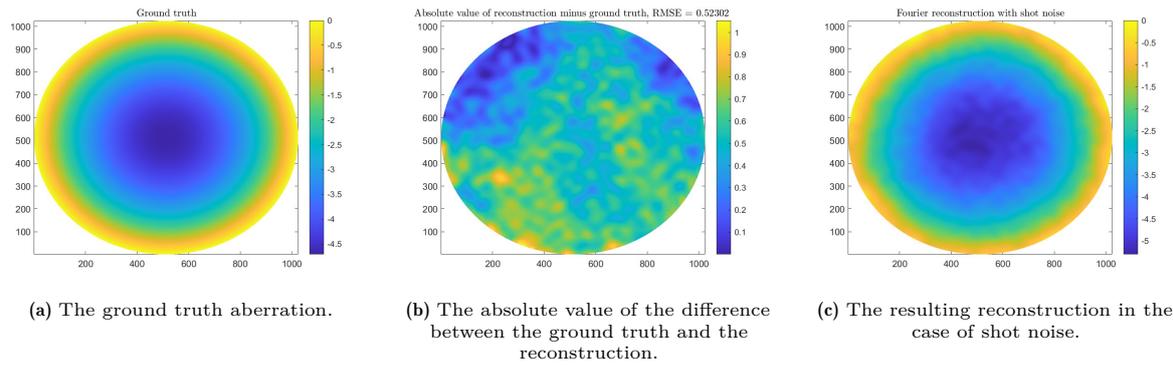


Figure 10.15: The ground truth, the reconstruction, and their absolute difference for the case of added shot noise on the Hartmanngram. Notice that despite the gradients being rather noisy, the algorithm can still reconstruct a wavefront that somewhat accurately describes the original.

As can be seen, the reconstructed wavefront still resembles the ground truth. This shows that the algorithm is rather robust when it comes to noise, as the gradients are clearly inconsistent with one another.

Alternative alignment

To show that the alternatively aligned Shack-Hartmann setup is also able to handle possible Poisson noise, the simulation is repeated for the alternative alignment. The alternatively aligned Shack-Hartmann sensor produces the following Hartmanngram:

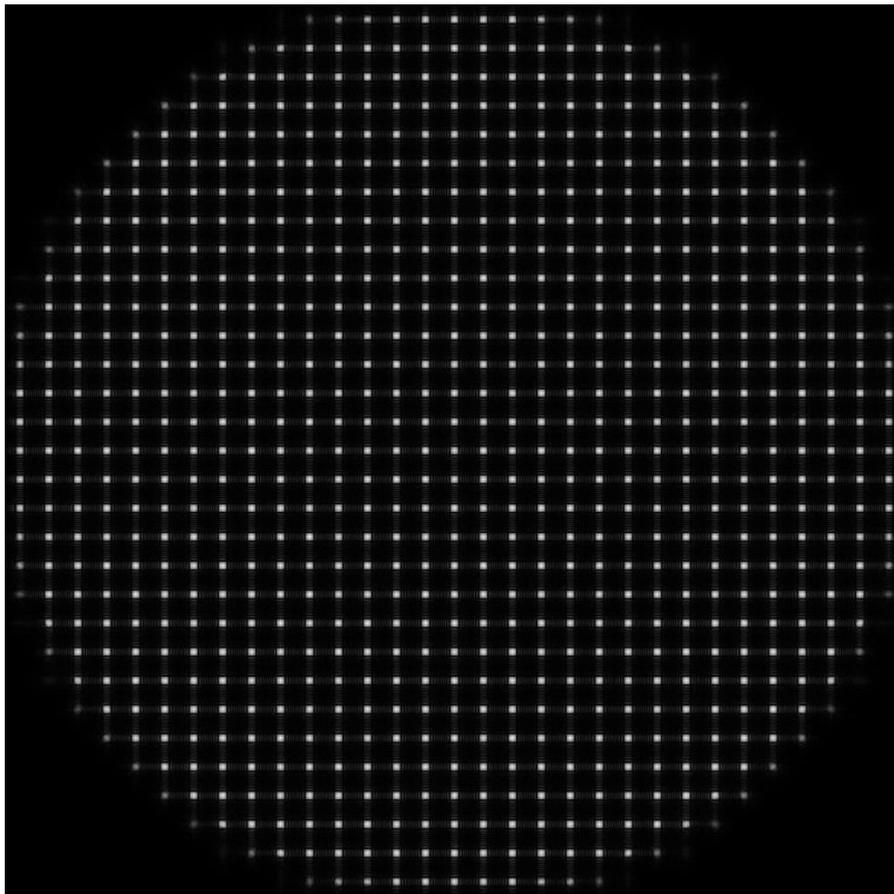
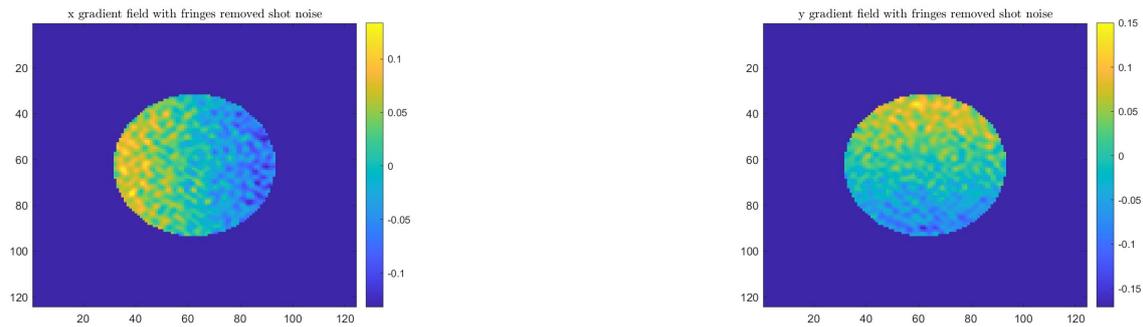


Figure 10.16: Alternative aligned Shack-Hartmanngram showing a defocus aberration with shot noise.

Calculating the gradients using the Fourier demodulation algorithm results in the following:

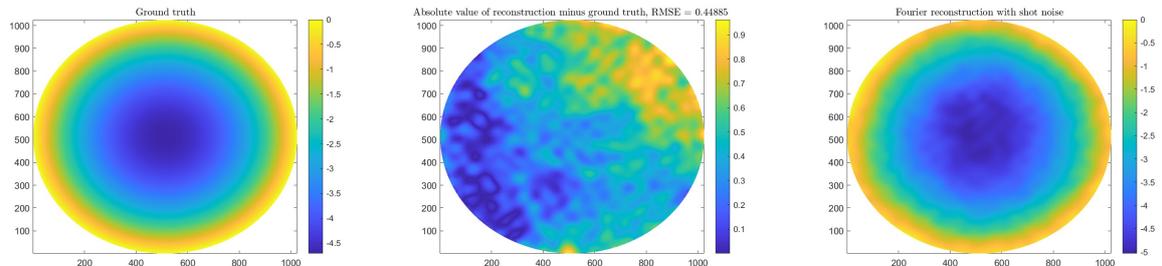


(a) The resulting x-gradient in the case of shot noise.

(b) The resulting y-gradient in the case of shot noise.

Figure 10.17: The resulting gradients in the case of shot noise on the Hartmanngram in the alternative alignment. Notice the noise is quite severe, these are gradients of a defocus aberration, but they are no longer nicely linear.

Again, the gradients are no longer linear and have been altered quite a bit. The resulting reconstruction, the ground truth and the difference between the two looks like this:



(a) The ground truth aberration.

(b) The absolute value of the difference between the ground truth and the reconstruction.

(c) The resulting reconstruction in the case of shot noise.

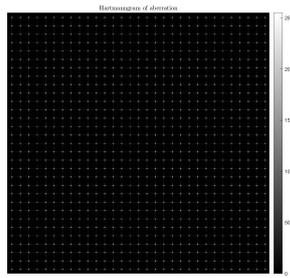
Figure 10.18: The ground truth, the reconstruction, and their absolute difference for the case of added shot noise on the Hartmanngram. Notice that despite the gradients being rather noisy, the algorithm can still reconstruct a wavefront that somewhat accurately describes the original.

10.5. Results using a square aperture

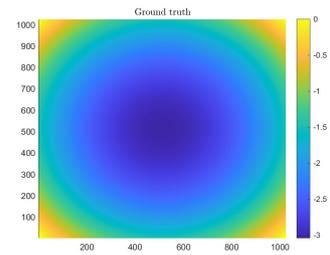
Hartmanngrams with square apertures are also analysed. As the circular aperture case already shows that the geometrical optics case performs a lot better than the physical optics, the physical optics Hartmanngrams are omitted.

10.5.1. Geometrical optics, traditional alignment

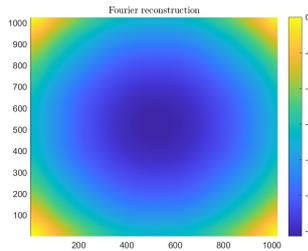
In this subsection the results for the traditional alignment using the geometrical Shack-Hartmann parameters are shown in a square aperture setting. The use of the geometrical parameters in the Shack-Hartmann sensor leads to a grid containing a high density of small spots without the effects of diffraction. These results can be compared against the ones in section 10.5.2 to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.2 to see the difference between the circular and square aperture.



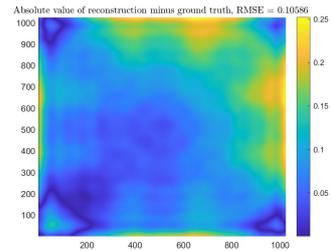
(a) Hartmanngram.



(b) Ground truth.

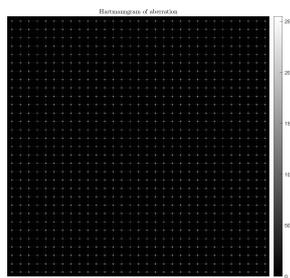


(c) Reconstruction using Fourier integration.

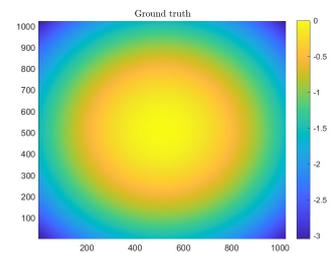


(d) Comparison between ground truth and reconstruction.

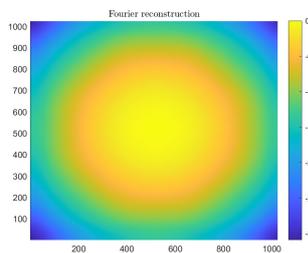
Figure 10.19: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a square aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.



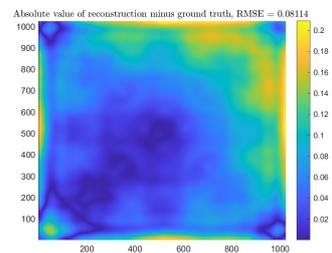
(a) Hartmanngram.



(b) Ground truth.

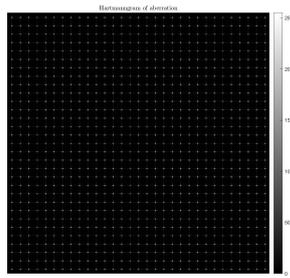


(c) Reconstruction using Fourier integration.

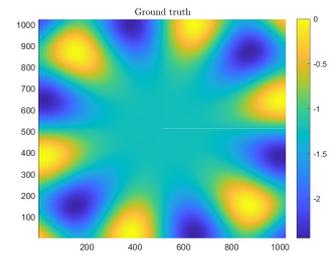


(d) Comparison between ground truth and reconstruction.

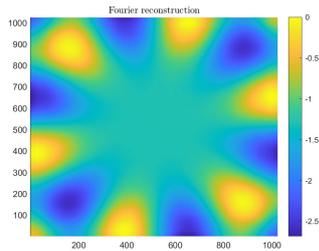
Figure 10.20: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a square aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{-3\pi}{2}$ is analysed.



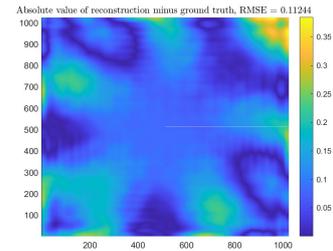
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.

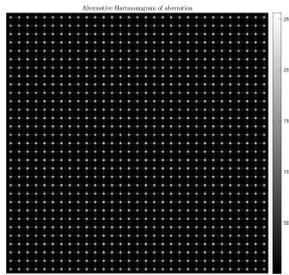


(d) Comparison between ground truth and reconstruction.

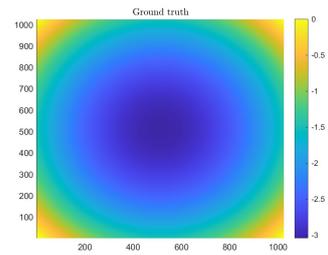
Figure 10.21: The Hartmanngram, reconstruction, ground truth and their differences for traditional alignment using a square aperture and a geometrical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

10.5.2. Geometrical optics, alternative alignment

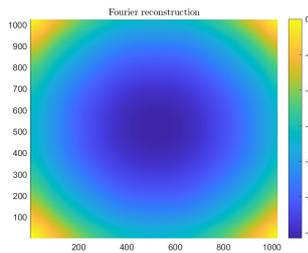
In this subsection the results for the alternative alignment using the geometrical Shack-Hartmann parameters are shown in a square aperture setting. The use of the geometrical parameters in the Shack-Hartmann sensor leads to a grid containing a high density of small spots without the effects of diffraction. These results can be compared against the ones in section 10.5.1 to see the difference between the alternative and the traditional alignment, and can also be compared against the results shown in section 10.2.1 to see the difference between the circular and square aperture.



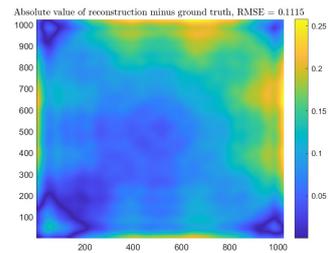
(a) Hartmanngram.



(b) Ground truth.

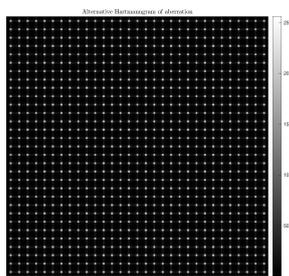


(c) Reconstruction using Fourier integration.

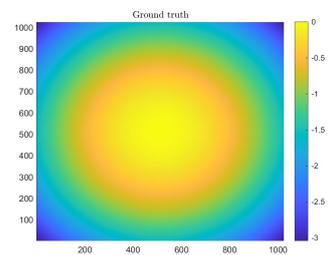


(d) Comparison between ground truth and reconstruction.

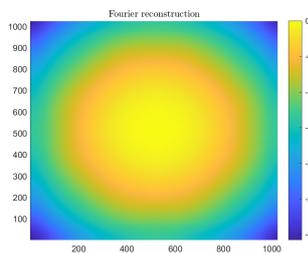
Figure 10.22: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a square aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{3\pi}{2}$ is analysed.



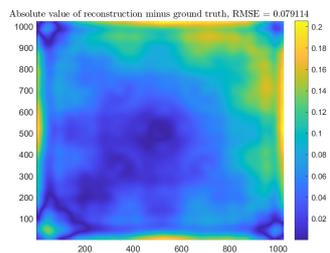
(a) Hartmanngram.



(b) Ground truth.



(c) Reconstruction using Fourier integration.



(d) Comparison between ground truth and reconstruction.

Figure 10.23: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a square aperture and a geometrical lenslet array. A defocus aberration of amplitude $\frac{-3\pi}{2}$ is analysed.

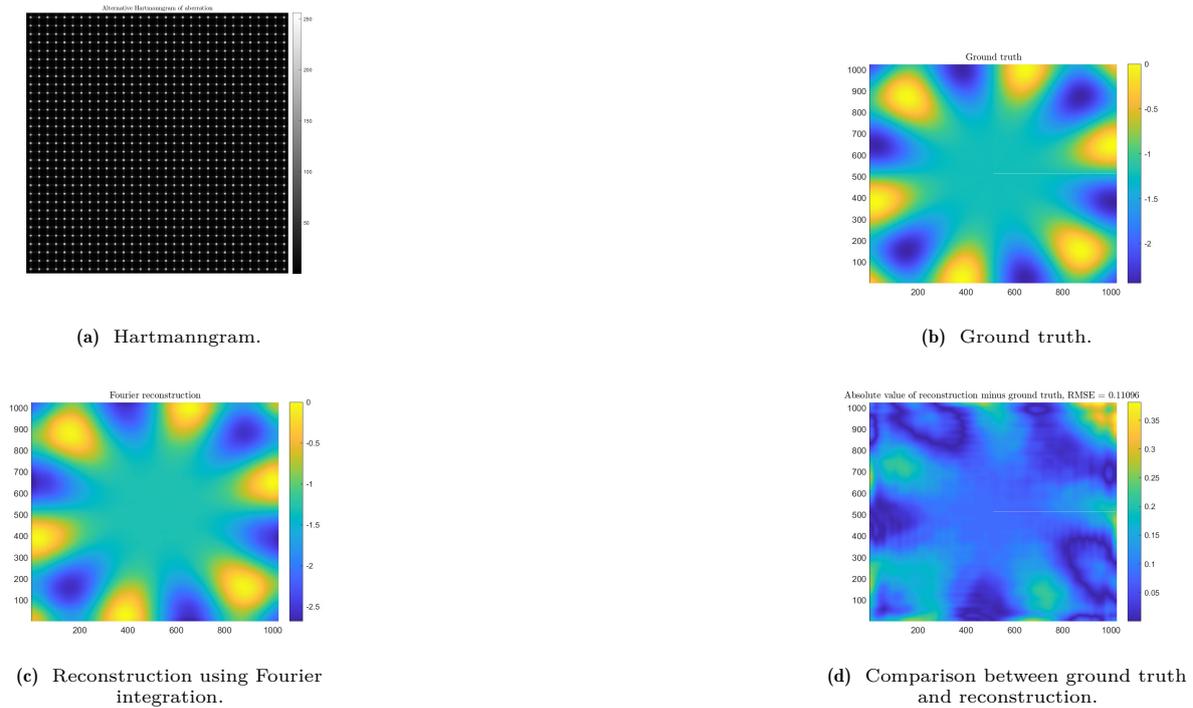


Figure 10.24: The Hartmanngram, reconstruction, ground truth and their differences for alternative alignment using a square aperture and a geometrical lenslet array. Zernike (14,6) was used as an aberration with amplitude $\frac{3\pi}{2}$ is analysed.

The resulting RMSE values are compared again:

	RMSE
Geometrical optics, traditional alignment	
Positive defocus	0.10586
Negative defocus	0.08114
High order Zernike (14,60)	0.11244
Geometrical optics, alternative alignment	
Positive defocus	0.1115
Negative defocus	0.079114
High order Zernike (14,60)	0.11096

Table 10.4: The different values of the RMSE shown in a table. Not that for positive defocus in a square aperture the traditional alignment actually outperforms the alternative alignment.

10.6. Discussion

Some things can be improved upon. First, the frontsurfer software is only able to fit the first 18 Zernike polynomials to the wavefront. This means that there can be no comparison between the Fourier demodulation and centroiding in the case of turbulence. It's RMSE value calculation is also flawed, and for some cases the Zernike coefficient differed so much from the ground truth that it was unclear what to make of it.

Second, for square apertures the Legendre polynomials can be used as an orthogonal basis function, just like the Zernike polynomials are used for circular apertures. Unfortunately Matlab does not have a Legendre function, whereas Zernike is provided. This would enable the comparison between square aperture wavefronts between the centroiding and the Fourier demodulation.

Real life data should of course be used to see if the method actually holds water in reality.

10.7. Summary

The combined use of the alternative alignment to use the Shack-Hartmann wavefront sensor as a curvature sensor and Fourier demodulation of to calculate the gradients of the wavefront is a viable way to measure the wavefront. It is shown that it works for both square and circular apertures, and also in the case of noise. It is shown that the alternative alignment slightly outperforms the traditional alignment in all but one scenario. It is also shown that the geometrical properties perform substantially better than the physical ones.

Conclusion of the thesis

11.0.1. Recalling the research question

The goal of this thesis is to answer the question: can alternative alignment and Fourier demodulation be used in combination with one another to reconstruct the wavefront. If this is the case, which Shack-Hartmann sensor properties are beneficial, and if there is an underlying difference, which aberrations are most suited to be analysed with this method. The answer will be twofold: theoretical work was done to gain insight in the way the two methods work on their own and also how they work together. Simulations were also run to quantify the performance.

11.1. Theoretical result

It was seen that the method of Fourier demodulation and the use of alternative alignment are both able to handle large amplitude aberrations.

Fourier demodulation does not require the spots to remain within a rigidly defined subaperture and is also able to handle the appearance and disappearance of spots in the aberration image compared to the reference image.

It was also seen that in a setup using traditional alignment the spots occupy a changing region when an aberration is introduced, and the amount of spots remains the same. In larger amplitude aberrations this means that the spots can occupy a region that becomes too large or too small. A setup using alternative alignment does not suffer from this problem: the region occupied by the spots on the CCD does not change, only the amount of spots changes. This means that larger amplitude aberrations should still be reconstructable using the alternative alignment.

Another advantage of the alternative alignment over the traditional alignment is that it is not necessary to know the spot shifts at the microlens array, instead they should be known at the CCD, which is how a Shack-Hartmann sensor operates. This means that the traditional alignment actually approximates the whole situation, whereas the alternative alignment precisely measures what is really happening.

11.2. Results from the simulations

It was seen that the alternative alignment paired with Fourier demodulation is a viable way of reconstructing the wavefront. The RMSE values are repeated here:

11.2.1. RMSE values for the square aperture

The table below shows the RMSE values retrieved from the reconstruction of various square fields:

	RMSE value
Geometrical optics, alternative alignment	0.20877
positive defocus	
Negative defocus	0.11166
Zernike (14,6)	0.14353
Geometrical optics, traditional alignment	0.21438
positive defocus	
Negative defocus	0.11329
Zernike (14,6)	0.14784
Physical optics, traditional alignment	0.8358
positive defocus	
Negative defocus	0.51652
Zernike (14,6)	0.66417
Physical optics, alternative alignment	0.68932
positive defocus	
Negative defocus	0.43357
Zernike (14,6)	0.61347

Table 11.1: RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.

11.2.2. RMSE values for the circular aperture

The table below shows the RMSE values retrieved from the reconstruction of various circular fields:

	RMSE value
Geometrical optics, alternative alignment	0.20877
positive defocus	
Negative defocus	0.11166
Zernike (14,6)	0.14353
Geometrical optics, traditional alignment	0.21438
positive defocus	
Negative defocus	0.11329
Zernike (14,6)	0.14784
Physical optics, traditional alignment	0.8358
positive defocus	
Negative defocus	0.51652
Zernike (14,6)	0.66417
Physical optics, alternative alignment	0.68932
positive defocus	
Negative defocus	0.43357
Zernike (14,6)	0.61347

Table 11.2: RMSE values compared for all twelve scenarios in the circular aperture. Two things are seen: the geometrical optics Shack-Hartmann sensor outperforms the physical optics setup, and second, the alternative alignment outperforms the traditional alignment.

11.2.3. Comparison with centroiding

Where it was possible, the performance of Fourier demodulation was compared to that of centroiding. The results can be seen in the table below.

	Ground truth coefficient	Ground truth residue	Fourier reconstruction coefficient	Fourier reconstruction residue	FrontSurfer coefficient	FrontSurfer residue
Circular 3π defocus geometrical traditional	2.3364	0.0225	2.3521	0.4604	2.5467	0.7607
Circular -3π defocus geometrical traditional	-2.3364	0.0225	-2.3849	0.4978	-2.5269	0.8644

Table 11.3: Table showcasing the difference in reconstruction coefficients for the Fourier demodulation method versus the centroiding method. Notice how the Fourier reconstruction coefficient is closer to the ground truth coefficient than the FrontSurfer coefficient, and how the residue is small for both circumstances.

11.3. Answer to the research question

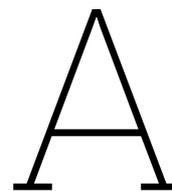
To answer the research question: yes, the combined use of alternative alignment with Fourier demodulation is recommended. It was shown that the reconstructions are more accurate than using traditional alignment with centroiding, alternative alignment with centroiding, and traditional alignment with Fourier demodulation. A Shack-Hartmann sensor with a low pitch and low focal length should be used, as Fourier demodulation gains no information from the diffraction patterns of the light, instead a grid of as many small spots should be used. As for the aberrations, there seemed to be no aberration in which the combined use of the alternative alignment with Fourier demodulation markedly outperformed the other methods. It seems the method is suited for both low and high order aberrations.

11.4. Advice for further research

For further research two main things should be done. First, the alternative alignment and Fourier demodulation should be combined in a real-life setup, such that real-life data is acquired on their combined use. Second, after having verified the system works in, the setup should be expanded to include a deformable mirror with which aberrations can be corrected for in real-time. This would show the performance in a closed loop setting. As a means to this the code used for Fourier demodulation should be optimised, as a lot can still be won in the speed of reconstructions.

References

- [1] A. Barbara. *Textbook on Keratoconus New Insights*. Jaypee Brothers Medical Publishers, 2011.
- [2] B. Boruah. *Wavefront sensing of light beams*. 2010. URL: <https://www.iitg.ac.in/physics/fac/brboruah/htmls/wfs.html> (visited on 06/27/2023).
- [3] J. Goodman. *Introduction to Fourier Optics*. 2nd ed. New York, America: McGraw-Hill, 1996.
- [4] E. Hecht. *Optics*. 5th ed. Essex, England: Pearson, 2017.
- [5] A. Kolmogorov. “Dissipation of energy in locally isotropic turbulence”. In: *Turbulence, Classic Papers on Statistical Theory* (1961).
- [6] D. Gavel & J. Brase L. Poyneer. “Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform”. In: *Journal of the Optical Society of America A* 19.10 (2002), pp. 2100–2111.
- [7] M. Verhaegen & G. Vdovin O. Soloviev. “Curvature sensing with a Shack-Hartmann sensor”. In: *Proceedings of the 10th International Workshop on Adaptive Optics for Industry and Medicine* (2005), pp. 1–4.
- [8] Shack & Platt. “Production and Use of a Lenticular Hartmann Screen”. In: *Journal of the Optical Society of America* 61.656 (1971).
- [9] H. Zebker & C. Werner R. Goldstein. “Satellite radar interferometry: Two-dimensional phase unwrapping”. In: *Radio Science* 23.4 (1988), pp. 713–720.
- [10] A. Talmi & E. Ribak. “Wavefront reconstruction from its gradients”. In: *Journal of the Optical Society of America A* 23.2 (2006), pp. 288–297.
- [11] Y. Carmon & E. Ribak. “Phase retrieval by demodulation of a Hartmann-Shack sensor”. In: *Optics Communications* 215 (2003), pp. 285–288.
- [12] F. Roddier. “Curvature sensing and compensation: a new concept in adaptive optics”. In: *Applied Optics* 27.7 (1988), pp. 1223–1225.
- [13] M. Roopashree, V. Akondi, and B. Prasad. “Performance analysis of Fourier and Vector Matrix Multiply methods for phase reconstruction from slope measurements”. In: 2009.
- [14] W.H. Southwell. “Wave-front estimation from wave-front slope measurements”. In: *Journal of the Optical Society of America* 70 (1980), pp. 998–1006.
- [15] S. Vallmitjana, A. Marzoa, and S. Bosch et al. “Testing two techniques for wavefront analysis. Specific applications and comparative study”. In: *Óptica Pura y Aplicada* 50.4 (2017), pp. 327–335.
- [16] M. Verhaegen, P. Pozzi, and O. Soloviev. *Control for High Resolution Imaging*. 2017.



Appendix

All the calculations in this thesis were performed by the author using a specially developed MATLAB package. The code contains general functions and algorithms. The code can be found in the "Control for scientific imaging systems - Delft center for systems and control" group on bitbucket. The link is <https://bitbucket.org/csi-dcsc/mscproject-seppe/src/master/>.